UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL

DNA BARCODING AND FORENSIC ENTOMOLOGY: A MOLECULAR APPROACH FOR DIPTERA SPECIES' IDENTIFICATION

Eva Sofia Alves Rolo

Mestrado em Biologia Humana e Ambiente

2010

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL

DNA BARCODING AND FORENSIC ENTOMOLOGY: A MOLECULAR APPROACH FOR DIPTERA SPECIES' IDENTIFICATION

Eva Sofia Alves Rolo

Mestrado em Biologia Humana e Ambiente

Dissertação orientada por:

- Professora Doutora Deodália Dias
Departamento de Biologia Animal & CESAM
Faculdade de Ciências da Universidade de Lisboa

- Professora Doutora Mª Teresa Rebelo
Departamento de Biologia Animal & CESAM
Faculdade de Ciências da Universidade de Lisboa

2010

FOREWORD

This study is part of the PTDC/SAU-ESA/101228/2008 project – Forensic Entomology: Morphometric and Molecular databank (mtDNA) to identify species (Diptera and Coleoptera) with forensic interest – funded by *Fundação para a Ciência e Tecnologia* (FCT).

This thesis was designed based on the preparation of two papers to be submitted to international journals. However, since this is an academic work (to get a master degree) it was considered important to devise a general introduction and a final consideration.

The articles are presented according to the standards of the journals for which these will be submitted:

- **Journal of Forensic Sciences** (American Academy of Forensic Sciences)

  - *Cytochrome* c *oxidase I effectiveness as a marker for insects' identification;*

  - *Forensic relevant insects' identification through GenBank and BOLD databases.*

AGRADECIMENTOS

O primeiro agradecimento é devido às minhas orientadoras. À Professora Deodália Dias pelo voto de confiança que depositou em mim desde o primeiro momento e pela capacidade de me fazer descontrair quando o stress aparece. À Professora Teresa Rebelo que despertou o meu interesse e curiosidade pelos insectos. Às duas agradeço todo o entusiasmo, o apoio e o incentivo que fizeram este trabalho andar para a frente.

À Joana Costa, o meu sincero e profundo "Muito Obrigado" porque foste quem me "abriu a porta" e me "guiaste" no laboratório e porque sem ti a análise deste trabalho simplesmente não existia. Obrigada pela paciência e dedicação. (Já tens o teu parágrafo!)

À Catarina, porque mais que colega de bancada foste um grande apoio durante todo este ano, e se aprendi muito, grande parte devo a ti.

Ao Eduardo e à Rita pelas sugestões e pela ajuda na correcção da dissertação. Vocês "deram cabo" de mim, mas as correcções foram essenciais.

Aos meus novos amigos, Catarina e Joana (outra vez), Ana Margarida, Sara, Ana Sofia, Sofia, Eduardo, Bruno, Francisco, Raquel, Lucas, Rui e Rita por todos os almocinhos, cafézinhos e lanchinhos que partilhámos, conversas sérias e conversas menos sérias, cinemas e problemas; enfim, por todos os momentos que passámos juntos. Vocês foram o bónus da minha tese. E que bónus! Obrigada.

À Vera, pelo que já partilhámos e pelo que ainda havemos de partilhar. Viva o Papagaio!!!

Aos meus outros amigos, por estarem lá sempre no final do dia.

À minha irmã Joana, porque tudo o que fazemos, fazemos juntas e, por isso, este trabalho também é teu. Amo-te muito Irmã!

Ao André, porque o teu dom é acreditares em mim quando eu não acredito, é orientares-me quando nada parece fazer sentido, é fazeres-me feliz

quando tudo parece contra… És muito importante para mim e sem ti não tinha sido possível. És assim simplesmente…

Por fim, queria agradecer às pessoas que me deram a maior "prenda" que alguma vez se pode querer. Sem vocês nunca seria quem sou hoje. Pai, Mãe, OBRIGADO de coração.

RESUMO

A Entomologia Forense é a ciência que aplica o conhecimento sobre os insectos, e outros artrópodes, em procedimentos jurídico-legais. O primeiro passo a ser tomado em Entomologia Forense é a identificação das espécies, normalmente realizada através de caracteres morfométricos, utilizando chaves dicotómicas de identificação; no entanto, a observação da morfologia é um método, por vezes, demorado e inconclusivo. Por outro lado, os métodos moleculares fornecem uma identificação rápida e precisa, possibilitam a identificação dos insectos em qualquer estádio de desenvolvimento, incluindo os estádios larvares, e podem ser utilizados independentemente das condições de preservação dos exemplares.

Na verdade, as metodologias para identificação molecular de espécies têm sofrido uma grande evolução e, actualmente, o *DNA barcoding* é considerado uma ferramenta muito útil na identificação de espécies. Este conceito baseia-se na amplificação e sequenciação de um pequeno segmento de DNA - conhecido como sequência *barcode* - de uma região padrão do genoma. Vários estudos sugerem o uso da sequência que codifica para a subunidade I da proteína citocromo *c* oxidase (COI) como o marcador de DNA adequado para o *DNA barcoding*. A identificação de espécies através desta nova ferramenta baseia-se na amplificação e sequenciação deste fragmento; uma vez obtida a informação da sequência do espécime-alvo é possível compará-la com sequências de referência, isto é, sequências de espécies previamente identificadas, já existentes numa biblioteca digital.

A identificação de espécies através do *DNA barcoding* implica, numa analise filogenética, que cada espécie surja como um grupo monofilético. Apesar, deste novo conceito se basear no uso de métodos de construção de árvores filogenéticas, não deve ser interpretado como tal, uma vez que a sequência *barcode* não apresenta, frequentemente, um sinal filogenético suficiente para determinar relações evolutivas. Um outro critério para a delineação de espécies

assenta em valores limite para as divergências nucleotídicas intra e interespecíficas. Um dos limites é de 3% (valor estabelecido para insectos), em que valores de divergência intra-específica abaixo deste limite determinam uma única espécie e valores de divergência interespecíficas acima, apontam para diferentes espécies. O outro limite, que surge como uma actualização do primeiro, sugere que a média da divergência nucleotídica entre espécies pertencentes ao mesmo género deve ser 10 vezes superior à média da divergência intra-específica encontrada para as mesmas espécies. A observação destes três critérios permite, assim, determinar se estamos perante a mesma espécie ou espécies diferentes

O *Barcode of Life Data System (BOLD)* é um *software* responsável pela gestão de dados obtidos através da ferramenta *DNA barcoding*. O sistema de identificação do *BOLD* é a unidade funcional para a identificação de espécimes no qual, a sequência obtida é submetida e comparada com as sequências referência, à semelhança dos sistemas utilizados noutros bancos de dados para a identificação de espécies (por exemplo, a base de dados GenBank do *National Center for Biotechnology Information,* NCBI).

A existência de evidências entomológicas pode ser de grande importância para casos forenses. De facto, estas podem fornecer informações importantes que poderão orientar o decorrer da investigação criminal.

A criação e implementação de uma Base de Dados de espécies de insectos é um passo importante para a Entomologia Forense. Com efeito, qualquer país que possua um serviço de Entomologia Forense eficaz e cientificamente bem suportado deve ter um conhecimento abrangente da diversidade de insectos. O uso do *DNA barcoding* sugere a sua utilidade na identificação de espécies de insectos encontrados em cenários forense. Apesar das vantagens científicas e pragmáticas existentes no conhecimento da diversidade de insectos em qualquer região do globo, a utilização deste marcador genético em bancos de dados exige que seja determinada a sua eficácia na distinção entre espécies.

Este estudo foi desenvolvido e escrito com vista à preparação de dois artigos científicos que serão submetidos a revistas internacionais da

especialidade. Neste sentido, a presente dissertação está dividida em quatro partes.

O Capítulo 1 refere-se à Introdução geral que assenta na revisão bibliográfica e estado de arte sobre a Entomologia Forense e do *DNA barcoding*, e que dá o fundamento ao trabalho desenvolvido.

O Capítulo 2 diz respeito ao primeiro artigo científico que tem como título "*Cytochrome* c *oxidase I effectiveness as a marker for insects' identification*". Este capítulo tem como principais objectivos determinar as sequências correspondentes à região do gene COI, de cada espécime, utilizada para efeitos de *DNA barcoding*, isto é, um fragmento de 658 pares de bases correspondente à região inicial do gene COI e, testar a eficácia deste para a identificação de espécies de dípteros com relevância forense. Aqui foram utilizados 52 indivíduos pertencentes a quatro espécies de Diptera, *Calliphora vicina* (Robineau-Desvoidy, 1830), *Calliphora vomitoria* (Linnaeus, 1758), *Lucilia caesar* (Linnaeus, 1758) e *Musca autumnalis* (De Geer, 1776). Estes espécimes foram recolhidos e morfologicamente identificados num estudo desenvolvido anteriormente. A amplificação, com *primers* universais, e a sequenciação da região em estudo foram facilmente obtidas. Este facto é muito vantajoso em situações que necessitam de uma maior rapidez na análise das amostras, como acontece em situações de contexto forense. O estudo filogenético permitiu identificar cada espécie como um grupo monofilético. Por sua vez, a análise das divergências nucleotídicas intra e interespecíficas, para as duas espécies do mesmo género, permitiram confirmar que, para os dois limites utilizados para a identificação de espécies através do DNA barcoding, estas são espécies diferentes. Estes resultados mostram a eficácia do COI como marcador genético para a discriminação de espécies.

O Capítulo 3 refere-se ao segundo artigo científico, e tem como título "*Forensic relevant insects' identification through GenBank and BOLD databases*". O principal objectivo deste trabalho foi determinar a capacidade destas bases de dados públicas para a identificação de espécies de insectos com interesse forense. Além disso, os dados foram também utilizados para determinar a eficácia do

marcador COI. Como anteriormente, todas as amostras foram facilmente amplificadas e sequenciadas. Os resultados mostraram que foi possível identificar 67.6% dos indivíduos, ao nível da espécie através da base de dados GenBank. Através da base de dados BOLD foi possível identificar 58.8% dos espécimes, também ao nível da espécie. No total foram identificados 49 espécimes pertencentes a 11 espécies diferentes: *Eudasyphora cyanella* (Meigen, 1826), *Lucilia caesar* (Linnaeus, 1758), *Pollenia rudis* (Fabricius, 1794), *Musca autumnalis* (De Geer, 1776), *Phaonia subventa* (Harris, 1780), *Phaonia tuguriorum* (Scopoli, 1763), *Helina impucta* (Fallén, 1825), *Helina evecta* (Harris, 1780), *Helina reversio* (Harris, 1780), *Hydrotaea dentipes* (Fabricius, 1805) e *Hydrotaea armipes* (Fallén, 1825). As sequências correspondentes a estas amostras foram utilizadas, posteriormente, para a análise filogenética e para o cálculo das divergências nucleotídicas intra e interespecíficas. Na análise filogenética foi possível observar situações de monofilia para todas as espécies. No que diz respeito à avaliação das divergências nucleotídicas entre espécies do mesmo género, os valores limite possibilitaram a discriminação de cada espécie. Em suma, estes resultados corroboraram a eficácia do gene COI para identificação de espécies.

Por fim, o Capítulo 4 destina-se às Considerações Finais, onde é referida a importância deste trabalho para a aplicação do marcador COI em bases de dados, utilizadas não só em situações de contexto forense mas também para o conhecimento global da diversidade biológica bem como a sua importância para a contribuição de uma base de dados da biodiversidade nacional.

Palavras-chave: *Entomologia Forense; DNA barcoding; Citocromo* c *Oxidase I; Diptera; Base de Dados; Barcode of Life Data System; GenBank.*

ABSTRACT

Forensic entomology is the science, which applies knowledge of insects (and other arthropods) to civil proceedings and criminal trials. Indeed, the existence of entomological evidences can be of great importance to forensic cases, because they can provide relevant information to delineate the course of the investigation; however, the species-level identification of specimens found on corpse is extremely important. Use of cytochrome *c* oxidase I (COI) as molecular marker for DNA barcoding project suggests that this approach could be very useful in forensic scene, where fast and accurate tools for species identification are essential. Molecular database implementation for insects' species is a very important step for the evolution of forensic entomology. Indeed, any country that wishes to have an effective and scientifically well supported forensic entomology service must have a comprehensive knowledge of insects' diversity.The main goals of this study are to provide evidence of the COI performance to be used as an effective, reliable and fast tool for an identification database and to determine what extent Barcode of Life Data System (BOLD) and GenBank databases are able, at that time, to identify insects' species with relevance. The COI fragment proposed for DNA barcode was sequenced and nucleotide sequence divergence within and between species and phylogenetic analysis were performed. In the two studies, COI allows observation of species discrimination as strongly supported monophyletic groups and intra and interspecific nucleotide divergences confirm the potential of COI in species delimitation. The results also showed that GenBank allowed to identify more sequences than BOLD, although the two databases have shown a good ability to identify insects' species.

Keywords: *Forensic Entomology; Cytochrome* c *Oxidase I; DNA barcoding; Database; Barcode of Life Data System; GenBank; Diptera.*

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AIC ..... Akaike information criterion
BOLD ..... Barcode of life data system
BOLD-IDS ..... Identification System of BOLD
bp ..... base pair
BSA ..... Bovine serum albumine
CBOL ..... Consortium for the Barcode of Life
COI ..... Citochrome c oxidase subunit I
DNA ..... Deoxyribonucleic acid
dNTP ..... Deoxyribonucleotide triphosphate
GTR+I+G ..... General time-reversible + Proportion Invariant + Gamma distribution shape parameter model
GTR+G ..... General time-reversible with gamma distribution shape parameter model
INML, IP ..... Instituto Nacional de Medicina Legal, Instituto Público
MCMC ..... Monte Carlo Markov Chain method
ML ..... Maximum Likelihood
MP ..... Maximum Parsimony
mtDNA ..... mitochondrial DNA
NCBI ..... National Center for Biotechnology Information
NJ ..... Neighbor-joining
NUMT ..... Nuclear mitochondrial pseudogene
PCR ..... Polimerase chain reaction
PMI ..... Postmortem interval
RNA ..... Ribonucleic acid
TBR ..... Tree Bisection and Reconnection algorithm

INDEX OF FIGURES

INDEX OF TABLES

Chapter 1

General Introduction

CHAPTER 1 – GENERAL INTRODUCTION

Entomology is derived from the Greek word *entomon* (insect) + *logos* (word, reason) meaning the study of insects (Gupta and Setia, 2004). Thus, forensic entomology is the science, which applies knowledge of insects (and other arthropods) to civil proceedings and criminal trials (Turchetto and Vanin, 2004).

According to Byrd (2006), forensic entomology commonly comprises three general areas: medicolegal or medicocriminal, urban, and stored product pests. The medicolegal area investigates the necrophagous feeding insects that colonise human corpses with legal purposes. The urban forensic entomology works with the insects that affect man and his immediate environment. Both the civil and criminal components of this area are involved, since the urban pests feed on both the living and the dead. Their mandibles can cause damages leading to economic problems. Besides, they can produce marks and wounds on the skin that may be misinterpreted as prior abuse. The stored products area deals with food and drink contamination by insects. The forensic entomology helps on determination of the insects' species involved, answers if their presence is accidental or intentional, and establishes if the levels of insects are allowable (Byrd, 2006). According to Anderson (1999), the wildlife forensic entomology should also be considered. This area assumes particular relevance in surveillance and protection of mistreatment of wild animals in captivity.

1.1 Retrospective

The first documented forensic entomology case is reported by the Chinese lawyer and death investigator Sung Tzu, in the 13th century. In his book, "Hsi yuan chi lu" (one possible translation is "The Washing Away of Wrongs") Sung

Tzu describes, possibly, the first case in which insects helped to solve a crime (Benecke, 2001; Amendt *et al.*, 2004; Gupta and Setia, 2004).

During medieval times, beyond the medical and legal experts, sculptors, painters and poets have closely observed the decomposition of human bodies and were made realistic and detailed illustrations of corpses containing maggots (Benecke, 2001) (Figure 1).

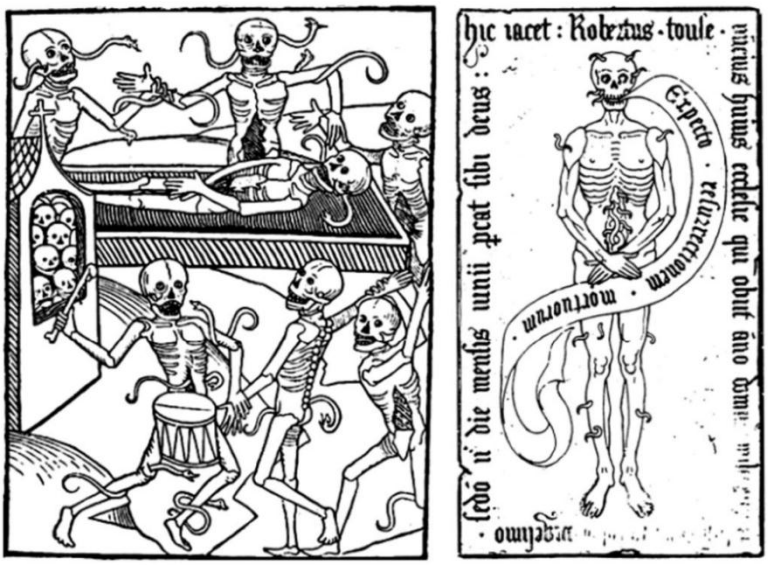

Figure 1. Illustration of corpses containing maggots: (left) "Dance of the Death" (15th century); (right) grave of Robert Touse (exact time of making unknown) (From: Benecke, 2001).

In 1855, the first modern forensic entomology case appeared, reported by Bergeret. He used forensic entomology to estimate the postmortem interval (PMI) (Benecke, 2001). Later, Yovanovich and Mégnin were the first forensic examiners who tried to evaluate insect succession on corpses, establishing properly the science of forensic entomology (Amendt *et al.*, 2004) and, in 1894, Mégnin published his most important book "La Faune des Cadavres", in which he explained his theory of eight successional insect's waves for freely exposed corpses (Benecke, 2001) and mentioned that on buried bodies insects came in two waves (Gupta and Setia, 2004). He also described the morphological features of various classes of insects that helped in their identification (Gupta and Setia, 2004).

Since the beginning of the 20th century, the interest in matter increased as well as the knowledge on the properties of insects. By now, forensic entomology has been accepted in many countries as an important tool and many studies have been made on the subject.

1.2 Postmortem changes of the human body

After death most animal bodies undergoes a process of decomposition which results in the gradual dissolution of the tissues (autolysis) into gases, liquids and salts caused essentially by proteolytic and other enzymes released by bacteria (Gordon *et al.*, 1988). Alternatively, an abnormal transformation of the corpse can occur depending on environmental conditions (maceration in immersed bodies, mummification in a dry environment) (Campobasso *et al.*, 2001).

During the decomposition, the body temperature decreases, phenomena known as *algor mortis*, and the skin color becomes red (*livor mortis* or lividity). Another sign of death is the stiffening of the muscle fibers due to the breakdown of glycogen and the accumulation of lactic acid (*rigor mortis*). Later skin slippage, the loosening of the epidermis from the underlying dermis occurs and hair and nails are easily detached. The production of a large quantity of gases during putrefaction causes physical distortion of the body, and a green coloration shows up the superficial blood vessels, the gastrointestinal region and those portions of the body where *livor mortis* was most marked. All these changes occur within the 72-96 h after death. Finally, when the temperature of the body is at the same level as the environment and following the initial putrefaction, no reliable estimation of the postmortem interval (PMI) is possible (Amendt *et al.*, 2004). Following this initial stage, also known as fresh stage, the body suffers others transformations according to more four main stages: putrefaction, dark putrefaction, butyric fermentation and dry stage (Bornemissza, 1957).

The postmortem decay rate can depend on intrinsic or extrinsic factors. The intrinsic factors comprise age and constitution of the body, cause of death, and integrity of corpse (Campobasso *et al.*, 2001). On the other hand, extrinsic factors like the ambient temperature, the humidity of the atmosphere, the movement of air or other medium, the state of hydration on the tissues, the nature of the medium, the nature of the soil and depth really influence the rate of decomposition (Gordon *et al.*, 1988). The existence of clothes can also slow down postmortem body cooling and favor the onset of the putrefaction process and also the animal predators, from arthropods to mammals, can have a predominant role in the breakdown of the corpse (Campobasso *et al.*, 2001).

1.3 Insects and the corpse

A cadaver constitutes a dynamic system that shelters and supports a rich community, of which arthropods form an important part, not only because they consume decomposing tissue but also because they speed up the decomposition processes (Arnaldos *et al.*, 2004). The colonization of a corpse by arthropods, and more precisely by insects, persists during the evolution of decomposition from the first few minutes after death until the bones resemble the bleached white stage (Haskell *et al.*, 1997).

1.3.1 Role of arthropods in decomposition

The cadaver can be colonized by a variable number of arthropods but only few species actively participate in cadaver breakdown directly accelerating the rate of decay (Campobasso *et al.*, 2001).

Each group of arthropods plays a different role in different stages of decomposition of organic matter. Its development in the cadaver is affected by several factors, temperature being the most important, affecting the rate of development and may cause diapause (the complete suspension of development)

(Myskowiak and Doums, 2002). Under favourable conditions, certain species of flies may lay their eggs or deposit larvae on exposed bodies. In the case of the egg-laying species, after a variable period, depending mainly upon the atmospheric temperature, the eggs hatch and the larvae feed upon the tissues, being loosed a considerable amount of tissue after death (Gordon *et al.*, 1988).

Colonizers species are selectively attracted by the decomposing status of the carrion. These species form complex communities within necrophagous species (also known as scavengers) which feed only on decomposing tissues, predators or parasites of the necrophagous species feeding on other insects or arthropods, omnivorous species feeding both on decomposing remains and associated arthropods, and other species which use the corpse as an extension of their habitat and part of their environment (Amendt *et al.*, 2004). In general, necrophagous, necrophilous and omnivorous are the most important groups in forensic studies. Within these, the necrophagous species that appear in a predictable sequence are the most important for forensic investigations (Arnaldos *et al.*, 2004).

1.3.2 Forensic evidence

The study of the order of appearance of arthropods on a corpse can provide conclusive evidence in a forensic case work (Arnaldos *et al.*, 2001). Indeed, the collection of arthropods found in a corpse has been shown to be very useful for estimating the time since death (Amendt *et al.*, 2000; Turchetto *et al.*, 2001; Wells *et al.*, 2001; Arnaldos *et al.*, 2004; Saigusa *et al.*, 2009).

According to Marchenko (2001), the scientific base of using entomological data in forensic entomology comprises: (1) existence of necrophagous insects in nature, which use cadaver tissues and pass the major part of their life cycle on cadavers; (2) relative constancy and specificity of cadaver entomofauna in a particular geographical region comprising widely spread predominating species; (3) compliance of species composition of cadaver entomofauna to the degree of its tissue decomposition and to its location; (4) seasonal alterations of predominant

necrophagous insect species; (5) beginning of insects activity in spring and its end in autumn as a result of transition to diapauses condition due to changes in temperature and light-time interval, the values thereof being dependent on geographical region and being specific for each species; (6) regulation of number of generations per vegetative period and of species life cycle duration by strictly definitive species-particular thermal parameters; (7) long preservation of insects chitin cuticles in nature.

1.3.3 Species with forensic relevance

For the purposes of forensic entomology, the two groups of insects most important are Diptera (flies) and Coleoptera (beetles) (Haskell *et al.*, 1997). Depending on the biogeographical region and ecological habitat, different species of insects are involved in the decay of a corpse; but generally, the first insects of the succession to colonize a cadaver belong to Diptera order.

In the Diptera, the blowflies species are the most important in forensic cases. These are the bright metallic blue and green "bottle" flies. Because of their huge number, the blowflies were the major vector in the degradation of the cadaver. They are mostly diurnal and usually rest at night (Chaubert *et al.*, 2003). Within the Diptera order, families like Calliphoridae, Sarcophagidae and Muscidae have a great relevance as forensic indicators (Arnaldos *et al.*, 2001). Calliphoridae and Muscidae were found to be the first to colonise the cadaver as soon as 2-3 h after exposure, followed by Sarcophagidae. The preferred oviposition sites were generally eyes, nasal openings, mouth, ears, and towards the end of the fresh stage the genitals (scrotum and vagina). According with the external temperatures hatching took place in a period ranging from 6 to 40 h after oviposition, larval development between 3 and 10 days and pupariation 6-18 days before emergence of adults. Fly activity continued until the dry stage of decomposition (Campobasso *et al.*, 2001).

The Coleoptera appearance increase both in number of species and in number of individuals in the later stages of body decomposition. Some Coleoptera

species colonize corpses as necrophagous insects while others are predators of Diptera larvae. Beetle activity (mainly Dermestidae) is essentially associated with the most advanced stages of the degradation process causing the drying out of semi-liquid soft tissues (Campobasso *et al.*, 2001). In case of Dermestidae, the larval stage, which are the real indicator of time since death, are characteristic of the most advanced stages of decomposition, even though adults specimens are known to appear in corpses from a very early time (Arnaldos *et al.*, 2004).

Other orders of insects known to frequent decomposing carrion include Hymenoptera (bees, wasps, ants), Lepidoptera (butterflies and moths), Hemiptera (true bugs), Dictyoptera (cockroaches), and Acari (mites) of the class Arachnida (spiders, ticks and mites). Of these groups, species of Hymenoptera are the most common. Wasps and ants are the main predators of fly eggs and larvae, while bees feed occasionally on fluids. Butterflies and moths have been observed to feed off of seepage from the carcasses, while bugs have been seen probing into the carrion, feeding in the underlying tissues. The cockroaches are usually found to cause superficial feeding artifacts on the surface of the skin of the corpse. They also may be liable for chewing off the eyebrows and eyelashes. In the order Acari, certain mite species are found to be associated with decomposing human remains. However, because they are very small, they are overlooked as evidence. These arthropods appear when remains are in advanced decay and drying, and they only are detected because they form aggregates and appear to be mold or piles of sawdust (Haskell *et al.*, 1997).

1.4 Importance of Forensic Entomology

Forensic entomology appears to provide answers to several questions that can be raised in a forensic case.

Firstly, forensic entomology intend to establish the time of death, known as postmortem interval (PMI), or more precisely, how long a carrion has been exposed in the environment. Indeed, using medical techniques, such as the

measurement of body temperature or analyzing livor and rigor mortis, time since death can only be accurately measured for the first 2 or 3 days after death. In contrast, by calculating the age of immature insect stages feeding on a corpse and analyzing the necrophagous species present on cadaver, postmortem intervals from the first day to several weeks can be estimated (Hall and Amendt, 2007). According to Hall and Haskell (cit by Haskell *et al.*, 1997), the PMI can be determinate using two entomological methodologies. The first is based on a known insect species life cycle (particularly the blowflies' life cycle) (Figure 2). The second method, proposed by Mégnin and others workers, is based on insect successional waves evaluations, that is, the nature of insect fauna present on the corpse at any given time (Figure 3).

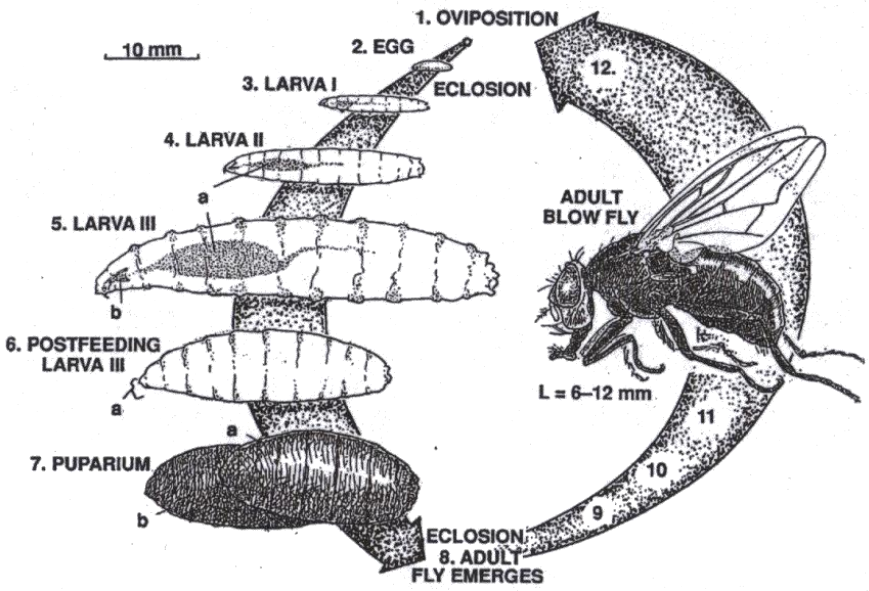

Figure 2. Example of a typical blowfly cycle. (1) Oviposition: eggs white to yellow. (2) Eclosion: maggot emerges. (3) Larva I: length about 10 mm. (4) Larva II: length 20 mm. (a) food in crop. (5) Larva III: length 45 mm, (a) blood in crop; (b) internal skeleton for feeding. (6) Postfeeding larva III: (a) internal feature obscured. (7) Puparium: changes color with age, (a) early stage; (b) late stage. (8) Eclosion: adult fly emerges. (9) After hardening, adult male and female flies seek mates. (10) Following copulation, female completes egg development. (11) Female lays egg mass (oviposits) on carrion/corpse at moist sites. (12) Female lay several egg masses in her adult life (1 to 3 weeks) (From: Haskell et al., 1997).

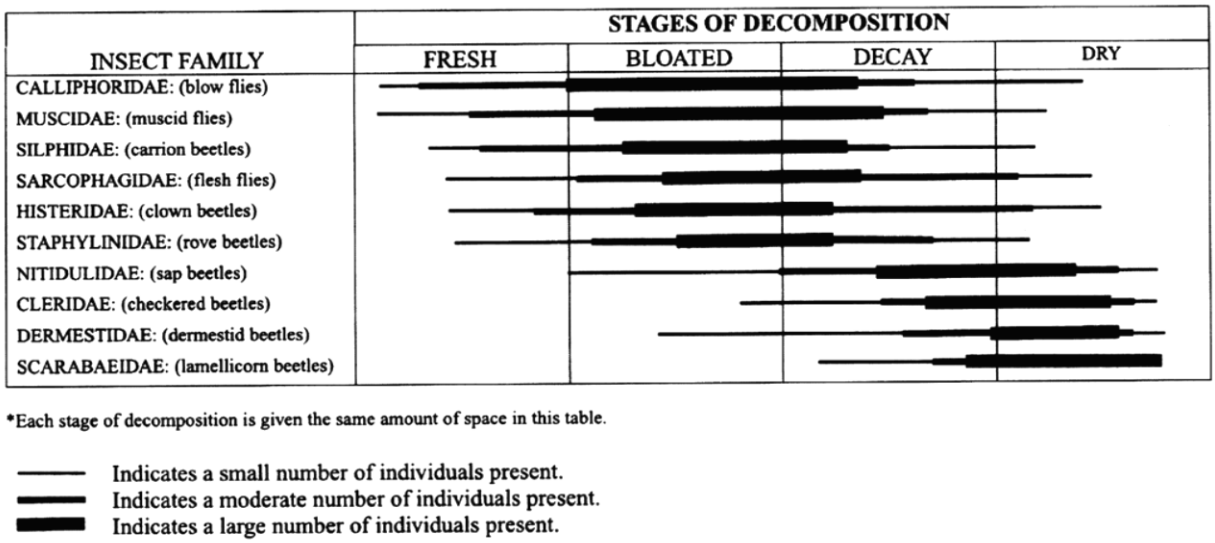

Figure 3. Example of adult arthropods succession on human cadavers in east Tennessee (during spring and summer) (From: Rodriguez and Bass, 1983).

Secondly, the ascertainment of postmortem transfer and, consequently, where was the initial location of the body, if it was hidden and where it was hidden can be made through the specimens' collection in the corpse. This is possible because, despite the fact that some common species are relatively ubiquitous, the presence of others species found only in certain geographical areas and occurred in a relatively definable environment (indoor or outdoor; rural and urban; wet or dry environment) can suggest that body was moved after death (Haskell *et al.*, 1997). Addicionaly, large accumulations of remnants (puparia of earlier generations of fly larvae, skins of beetle larvae, the bodies of dead insects and larvae solid excrements) left by insects occur when a decomposing body lies for a long period, and this can help to confirm that the body has lain undisturbed *in situ* for an extended time (Archer *et al.*, 2005). In the same way, the presence of live maggots or remnants of insects in the absence of a dead body at a location is almost certain evidence that some kind of corpse has been removed from the scene (Campobasso and Introna, 2001).

Forensic entomology is also used in diagnosis of poisoning. Indeed, when bodies are in a state of advanced decomposition or that are skeletonized the examination for toxicologically relevant substances may be difficult due to the

lack of appropriate sources such as tissue, blood or urine (Amendt *et al.*, 2004). On the other hand, maggots feeding on intoxicated tissues introduce into their own metabolism drugs and toxins (Campobasso and Introna, 2001) that will be deposited into fat bodies and the exoskeletal material (chitin) of the insect. These ingested drugs are sheltered into the chitin and remain in the specimen for an extended period of time (Haskell *et al.*, 1997). Consequently, a thorough toxicological analysis of necrophagous larvae and remains from a corpse may be crucial to the correct determination of death (Campobasso and Introna, 2001). However, it is known that toxics modified the development rate of maggots and the use of insect life stage method in calculation of PMI must be careful to avoid errors in PMI estimation.

Other aim of forensic entomology is the detection of negligence situations (Benecke and Lessig, 2001; Anderson and Huitson, 2004; Archer *et al.*, 2005). The early colonization of living people and animals is known as myasis, and the occurrence of maggots in wounds or natural orifices may indicate negligence and can help to estimate how long this situation of neglect was verified. Although this advantage, these colonizers are of the same species found in early decomposition stage of corpses and this can lead to complications in estimation of PMI.

Finally, other questions like the time of decapitation and/or dismemberment, the submersion interval, the identification of specific sites of injury on the body and postmortem artifacts (both, on the body and in the crime scene), the suspect association to crime scene, and sexual molestation can be answered through entomological investigation.

These findings can then inform several stages of the criminal justice process: the initial scene investigation, the subsequent follow-up investigative process when evaluating suspects and witnesses, and the criminal trial.

1.5 DNA barcoding

Accurate identification of an insect specimen is usually a crucial first step in a forensic entomological analysis. Closely related carrion species can substantially differ in growth rate, diapause response or ecological preferences. Species-diagnostic based on anatomical characters are not known for the immature stages of many forensically important insects and an existing key may be incomplete or difficult for non-specialists to use (Wells and Stevens, 2008), and the correct species determination is indispensable in forensic investigations.

The identification of insects based on deoxyribonucleic acid (DNA) can be performed with immature insects or fragments of puparium and adult insects, and provide a much faster identification and thus facilitate the successful conclusion of a case (Harvey *et al.*, 2003; Mazzanti *et al.*, 2010). According to Amendt *et al.* (2004) polimerase chain reaction (PCR) amplification of suitable regions of the genome, sequence analysis of the amplicons obtained, and alignment of the data with reference sequences is the usual and recommended method.

Today, the concept of DNA barcoding arises as a molecular approach to identify species. This concept is based on a DNA sequence that acts as a barcode specific for each species (Hebert *et al.*, 2003). In this way, the DNA barcode is a short sequence of nucleotides taken from an appropriate part of an organism's genome that is used to identify it at species level.

Species identification by DNA barcoding is a sequencing-based technology. Once obtained the sequence information of the target specimen it is possible comparing this information to a sequence library from known species (Hajibabaei *et al.*, 2007). Nowadays, several libraries of DNA sequences can be found. Some of these repositories are comprehensive and include sequences from several segments of DNA (e.g. GenBank), but others are restricted to a specific marker (e.g. BOLD) (see Chapter 3).

The key point for any taxonomic system is its ability to deliver accurate species identification and, according to Hebert *et al.* (2003), DNA barcoding accurately identified species in more than 95% of cases.

1.5.1 Nuclear DNA versus Mitochondrial DNA

Generally, the mitochondrial genome (mtDNA) of animals is a better target for analysis than the nuclear genome because of its high copy number, lack of introns, its limited exposure to recombination and its haploid mode of inheritance (Hebert *et al.*, 2003) and therefore, have an increased chance of generating species-specific markers (Harvey *et al.*, 2003). In animals, mtDNA occurs as a single double-helical circular molecule containing 13 protein-coding genes, 2 ribosomal genes, a non-protein coding control region, and several transference RNAs. Each mitochondrion contains several such circular molecules and, therefore, several complete sets of mitochondrial genes. Furthermore, each cell has several mitochondria. Thus, when sample tissue is limited, the mitochondrion offers a relatively abundant source of DNA (Waugh, 2007). Consequently, these features make the mtDNA clearly advantageous to forensic studies where material may be only fragments or poorly preserved.

1.5.2 Cytochrome c oxidase subunit I (COI) as DNA barcoding marker

The efficacy of DNA barcoding depends on selection of a suitable segment of DNA. Indeed, its mutation rate must be slow enough so that intraspecific variation is minimised but sufficiently rapid to highlight interspecific variation, it must be relatively easy to collect, and should have as few insertions or deletions as possible to facilitate sequence alignment (Hebert *et al.*, 2003).

In 2003, Herbert *et al.* published a study in which they suggest the use of cytochrome *c* oxidase I as the suitable DNA marker to DNA barcoding.

Eukaryotic cytochrome *c* oxidase, the last enzyme of the mitochondrial respiratory chain, is highly conserved across species that employ oxidative

phosphorylation for metabolism and is a multimeric enzyme of dual genetic origin. The subunits I, II and III are large transmembrane proteins, highly hydrophobic, encoded in mitochondrial genome (Figure 4). The remaining small subunits that surround the core of the enzyme are encoded in the nuclear genome (Fontanesi *et al.*, 2008). Cytochrome *c* oxidase subunit I (COI), the catalytic subunit of the enzyme, is predominantly imbedded in the membrane of the mitochondrial crista. The nucleotides of the gene that codes for it show sufficient variation to differentiate between species (Waugh, 2007). Indeed, Hebert *et al.* (2003) says that COI have two important advantages: (1) the universal primers for this gene are very robust, enabling recovery of its 5' end from representatives of most, if not all, animal phyla and (2) COI appears to possess a greater range of phylogenetic signal than any other mitochondrial gene (the evolution of this gene is rapid enough to allow the discrimination of not only closely allied species, but also phylogeographic groups within a single species).

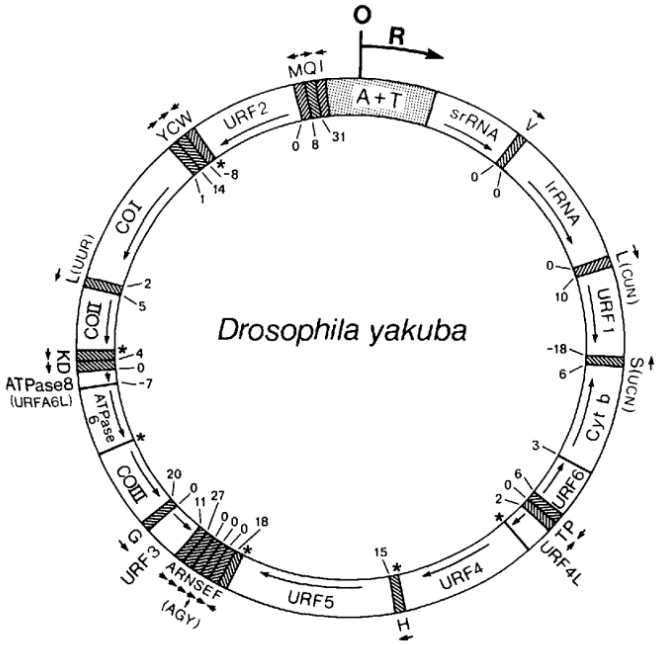

Figure 4. Gene map of the *D. yakuba* mtDNA molecule (From: Clary and Wolstenholme, 1985).

However, according to Frézal and Leblois (2008), the DNA barcoding shows some crucial pitfalls. First, the existence of under-described fraction of

biodiversity complicate the identification of unknown specimens, since the individuals chosen to represent each taxon in the reference database could not cover all of existing diversity in this taxon. Second, the inherent risks due to mitochondrial inheritance can lead to over- or underestimate sample divergence and render conclusions on species status unclear. Indeed, heteroplasmy (i.e. the presence of a mixture of more than one type of mitochondrial genome within a single individual), and maternally transmitted bacteria (e.g. Wolbachia, Whitworth *et al.*, 2007) can cause misleading processes in identification. Third, nuclear mitochondrial pseudogenes (NUMTs), this is non-functional copies of mitochondrial DNA sequences translocated into the nuclear genome (Song *et al.*, 2008), could mimic mitochondrial copies of COI introducing ambiguity into the barcoding and lead to disturbances in specimens' identification. Fourth, the rate of evolution in COI marker, since the evolution rate is not equal for all living species, can lead to a lack of resolving power. Finally, the intra-specific geographical structure can generate high rates of intra-specific divergence that can blur and distort species delineation.

Despite these shortcomings, DNA barcoding may prove to be an efficient tool for rapid assessment of taxonomic diversity, especially in species groups that are otherwise difficult to study (Linares *et al.*, 2009) and, consequently, could be very helpful in forensic entomology investigations (see Chapter 2).

1.6 Framing in Master degree

The difficulties in morphological identification of some insects and the possible association of these to a forensic context show the necessity of molecular identification of species found in these scenarios.

The content of this dissertation intents to understand the importance of Forensic Biology, both in the areas of Molecular Biology and Genetics, and in Forensic Entomology either when applied to legal and criminal research.

Under the master's degree in *Biologia Humana e Ambiente*, this work comes as a contribution to cover the gap in forensic entomology in Portugal, particularly in the molecular systematic characteristic of insects. Moreover this will be the first step in the creation of the National Molecular Database of insects' species with forensic relevance based on a new concept for species identification, the DNA barcoding.

1.7 Main goals

In Portugal, forensic entomology is still a very undeveloped area and this study appears to cover this gap.

For purposes of this study, we will focus our attention in medicolegal and wildlife forensic entomology, because the involvement of insects in decomposition of cadavers.

Thereby, the main goals of this study are:

- determine the DNA barcoding sequences of some insects' species (previously identified by morphological methods);
- test the effectiveness of the COI for the identification;
- evaluate if the databases that currently exist (e.g. GenBank from NCBI; BOLD from CBOL) are able to identify species with forensic relevance based on COI sequence;
- contribute to the implementation of a National Molecular Database applicable for Portugal area.

Despite these main objectives, this thesis aim the acquirement of qualification in laboratorial practice and in analysis of the results obtained during the laboratorial work.

References

Amendt J., Krettek R., Niess C., Zehner R., Bratzke H. 2000. Forensic entomology in Germany, Forensic Sci. Int. 113: 309-314.

Amendt J., Krettek R., Zehner R. 2004. Forensic entomology. Naturwissenschaften 91: 51-65.

Anderson G.S. 1999. The use of insects to determine time of death in illegally killed black bears. J Forensic Sci. 44(4): 843-846.

Anderson G.S., Huitson N.R. 2004. Myiasis in pet animals in British Columbia: The potential of forensic entomology for determining duration of possible neglect. Can Vet J. 45: 993-998.

Archer M.S., Bassed R.B., Briggs C.A., Lynch M.J. 2005. Social isolation and delayed discovery of bodies in houses: the value of forensic pathology, anthropology, odontology and entomology in the medico-legal investigation. Forensic Sci. Int. 151: 259-265.

Arnaldos I., Romera E., Garcia M.D., Luna A. 2001. An initial study on the succession of sarcosaprophagous Diptera (Insecta) on carrion in the southeastern Iberian peninsula. Int. J. Legal Med. 114: 156-162.

Arnaldos M.I., Sánchez F., Álvarez P., García M.D. 2004. A forensic entomology case from the Southeastern Iberian Peninsula. Aggrawal's Internet J. Forensic Med. Toxicol. 5(1): 22-25.

Benecke M. 2001. A brief history of forensic entomology. Forensic Sci. Int. 120: 2-14.

Benecke M., Lessig R. 2001. Child neglect and forensic entomology. Forensic Sci. Int. 120:155-159.

Bornemissza G.F. 1957. An analysis of Arthropod succession in Carrion and the effect of its decomposition on the soil fauna. Aust J Zool. 5(1): 1-12.

Byrd J.H. 2006. Forensic Entomology [www.forensiccareers.com/index.php?option=com_content&task=view&id=21&Itemid=24 – last access 4.01.2009].

Campobasso C.P., Introna F. 2001. The forensic entomologist in the context of the forensic pathologist's role. Forensic Sci. Int. 120: 132-139.

Campobasso C.P., Vella G.D., Introna F. 2001. Factores affecting decomposition and Diptera colonization. Forensic Sci. Int. 120: 18-27.

Chaubert S., Wyss C., Cherix D. 2003. Suicide in a forest – Determining post-mortem interval with blowfly and fleshfly species (Diptera: Calliphoridae, Sarcophagidae). Forensic Sci. Int. 136(Supplement 1): 388.

Clary D.O., Wolstenholme D.R. 1985. The Mitochondrial DNA Molecule of *Drosophila yakuba*: Nucleotide Sequence, Gene Organization, and Genetic Code. J Mol Evol. 22: 252-271.

Fontanesi F., Soto I.C., Barrientos A. 2008. Cytochrome *c* Oxidase Biogenesis: New levels of Regulation. IUBMB Life. 60(9): 557-568.

Frézal L., Leblois R. 2008. Four years of DNA barcoding: Current advances and prospects. Infect Genet Evol. 8: 727-736.

Gordon I., Shapiro H.A., Berson S.D. 1988. Diagnosis and the early signs of death, in: I. Gordon, H.A. Shapiro, S.D. Berson (eds.), Forensic Medicine: A Guide to Principles, Churchill Livingstone, Edinburgh and New York, pp. 43-58.

Gupta A., Setia P. 2004. Forensic entomology – past, present and future. Aggrawal's Internet J. Forensic Med. Toxicol. 5(1): 50-53.

Hajibabaei M.,. Singer G.A.C., Clare E.L., Hebert P.D.N. 2007. Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. BMC Biol. 5: 1-24.

Hall M., Amendt J. 2007. Forensic entomology – scientific foundations and applications. Forensic Sci. Int. 169S: S27-S28.

Harvey M., Dadour R.I., Gaudieri S., Mansell M.W., Villet M.H. 2003. MtDNA-based identification of blowflies in forensic entomology: more accurate postmortem interval estimation and beyond. Forensic Sci. Int. 136(Supplement 1): 389.

Haskell, N.H., Hall R.D., Cervenka V.J., Clark M.A. 1997. On the body: Insects' life stage presence and their postmortem artifacts, in: Forensic Taphonomy - The Postmortem Fate of Human Remains, W.H. Haglund, and M.H. Sorg, Eds., CRC Press, Boca Raton, FL, pp. 415-448.

Hebert P.D.N., Cywinska A., Ball S.L., deWaard J.R. 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B. 270: 313-321.

Linares M.C., Soto-Calderón I.D., Lees D.C., Anthony N.M. 2009. High mitochondrial diversity in geographically widespread butterflies of Madagascar: A test of the DNA barcoding approach. Mol Phylogenet Evol. 50: 485-495.

Marchenko M.I. 2001. Medicolegal relevance of cadaver entomofauna for the determination of the time of death. Forensic Sci. Int. 120: 89-109.

Manzatti M., Alessandrini F., Tagliabracci A., Wells J.D., Campobasso C.P. 2010. DNA degradation and genetic analysis of empty puparia: Genetic identification limits in forensic entomology. Forensic Sci Int. 195: 99-102.

Myskowiak J.B., Doums C. 2002. Effects of refrigeration on the biometry and development of *Protophormia terraenovae* (Robineau–Desvoidy) (Diptera: Calliphoridae) and its consequences in estimating post-mortem interval in forensic investigations. Forensic Sci Int. 125: 254-261.

Rodriguez W.C., Bass W.M. 1983. Insect activity and its relationship to decay rates of human cadavers in East Tennessee. J Forensic Sci. 28:423–432.

Saigusa K., Matsumasa M., Yashima Y., Takamiya M., Aoki Y. 2009. Practical applications of molecular biological species identification of forensically important flies. Legal Med. 11: S344-S347.

Song H., Buhay J.E., Whiting M.F., Crandall K.A. 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc Natl Acad Sci. 105(36): 13486-13491.

Turchetto M., Vanin S. 2004. Forensic entomology and globalization. Parassitologia 46(1-2): 187-190.

Turchetto M., Lafisca S., Costantini G. 2001. Postmortem interval (PMI) determined by study sarcophagous biocenoses: three cases from the province of Venice (Italy). Forensic Sci Int. 120: 28-31.

Waugh J. 2007. DNA barcoding in animal species: progress, potential and pitfalls. Bioessays 29: 188-197.

Wells J.D., Stevens J.R. 2008. Application of DNA-based methods in forensic entomology. Annu. Rev. Entomol. 53: 103-120.

Wells J.D., Introna F., Jr, Di Vella G., Campobasso C.P., Hayes J., Sperling F.A.H. 2001. Human and insect mitochondrial DNA.analysis from maggots. J Forensic Sci. 46(3): 685–687.

Whitworth T.L., Dawson R.D., Magalon H., Baudry E. 2007. DNA barcoding cannot reliably identify species of the blowfly genus *Protocalliphora* (Diptera: Calliphoridae). Proc R Soc B. 274: 1731-1739.

Chapter 2

Cytochrome c *oxidase I effectiveness as a marker for insects' identification*

CHAPTER 2 – CYTOCHROME *C* OXIDASE I EFFECTIVENESS AS A MARKER FOR INSECTS' IDENTIFICATION

Abstract

The implementation of a molecular database of insects' species is very important step for the evolution of forensic entomology. Indeed, any country that wishes to have an effective and scientifically well supported forensic entomology service must have a comprehensive knowledge of insects' diversity.

The widespread use of cytochrome *c* oxidase I (COI) as the ideal molecular marker for DNA barcoding project suggests that this approach could be very useful as well in forensic scene, where rapid, precise species identification tools are vital. Despite scientific and pragmatic advantages of knowing the diversity of insects with forensic interest through the globe, the implementation of such molecular database requires the establishment of its ability to distinguish different species in forensics too.

Using four common fly species found to be forensically relevant (*Calliphora vicina*, *Calliphora vomitoria*, *Lucilia caesar* and *Musca autumnalis)*, this study aimed to provide evidence of the COI performance to be used as an effective, reliable and fast tool for an identification database.

The COI fragment proposed for DNA barcode was sequenced; then, nucleotide sequence divergence within and between species and phylogenetic analysis were performed.

Phylogenetic analyses show all species as strongly supported monophyletic groups. The intraspecific divergence within *Calliphora* shows an average value of 0.24% and average of interspecific divergence percentage between these congeneric species was 4.9%. Highest interspecific divergence values occur between *M. autumnalis* and the other three species. In fact, this

species belongs to Mucidae while other three belongs to Calliphoridae, being phylogenetically more distant.

According to our molecular data, this method appears to be an accurate and robust technique for identifying at least these most common fly species with forensic relevance.

Keywords: forensic science; database; forensic entomology; Diptera; cytochrome *c* oxidase I; DNA barcoding.

1. Introduction

DNA barcoding is a new molecular tool useful in species discrimination, which uses a small DNA fragment – known as DNA barcode – from a standardized region of the genome (1). This fragment consists of a 658 bp string corresponding to nucleotide positions 1490-2198 from the 5'– end of cytochrome c oxidase subunit I gene (COI) using *Drosophila yakuba* mitochondrial genome as a reference (2).

Forensic entomology studies the interaction of insects and other arthropods with dead bodies, and like other forensic sciences, is used for legal purposes (3). Different insect species colonizing corpses have different biologies (life-cycle, ecological preferences, distribution, etc.) and, based on this, a forensic entomologist can provide answers for several questions in a crime scene: estimation of postmortem interval (PMI), postmortem transfer, diagnosis of poisoning, and neglect of living people (4). Since corpses' colonization occurs by successive waves and colonization pattern changes regionally and seasonally, identifying which species colonize the corpse is the key for forensic entomologist work. Thus, identification of insects collected from a corpse must be precise; otherwise, erroneous developmental data application may result in an incorrect PMI estimation (3).

In this respect, species have been widely identified through the use of mostly morphological criteria. Morphological identification consists in anatomical character-based keys, only usable by few experts, to identify the adults (or larvae and pupae in some cases) to species level. However, for most groups, keys when available can be vague and the identification can become difficult and almost impossible. In addition, larval stage is the most usually found on corpses (5) and time-consuming rearing of this stage to adult for identification may delay criminal investigation or cause significant problems when rearing fails (6). Under these circumstances, species' identification based on molecular analysis can appears as a more suitable way for unknown specimens' identification. Compared with morphological identification, molecular data acquisition arises as a less time consuming methodology and can also be the only way to identify damaged organisms or fragments, very common in forensic scenarios (7,8). Furthermore, molecular identification can be the only way when there are no obvious means to match adults with immatures, and when morphological traits do not clearly discriminate species (9).

Using DNA barcoding concept for insects' species identification should be taken into account three main criteria for species delimitation:

1) The use of a threshold value, to separate intraspecific from interspecific variation, the so-called "barcoding gap" (1,10). For example, in insects, genetic distance between different species almost always exceeds 3% (1);
2) The second criterion comes as an update of the previous, and suggests that this threshold value should be ten times greater than the average of intraspecific nucleotide distance for different animal species (11);
3) Finally, the monophyletic association of specimen within a species in a phylogenetic analysis is required for a successful species' identification (12,13), that meaning each morphological species should appears in a single monophyletic lineage (14). However, in spite of this method uses a phylogenetic tree construction method, this should not be interpreted as phylogenies, since DNA barcodes do not frequently demonstrate

sufficient phylogenetic signal to determine evolutionary relationships (15).

After sequencing, an unknown insect sequence can be compared with a library of barcode reference sequences obtained from specimens of known identity. If it matches with a high confidence level with a reference sequence, it can be assumed that the unknown specimen belongs to the reference taxon (species) or, at least, to the group with identical species. on the other hand, if the unknown sequence does not match with any within the database, new data can be recorded as a new haplotype or a geographical variant, or can suppose the unveiling of a new species (6,15). Finally, information can be crossed with prior knowledge regarding developmental stages of each species and ecological data, and allows determination of relevant aspects with medicolegal purposes, including PMI.

However, before assuming the use of COI as a molecular tool in forensic entomology, it's necessary to ascertain their suitability on insects' species identification. In this way, several specimens of Diptera, *Calliphora vicina* (Robineau-Desvoidy, 1830), *Calliphora vomitoria* (Linnaeus, 1758), *Lucilia caesar* (Linnaeus, 1758), all belonging to Calliphoridae, and *Musca autumnalis* (De Geer, 1776), belonging to Muscidae, were sequenced with the intent to evaluate COI effectiveness for implementation of this DNA barcoding marker in databases for the identification of insect species with forensic interest.

2. Materials and Methods

2.1 Samples

Insect specimens used in this work were obtained in a previous study (16). Samples were collected from mammalian carcasses air exposed, in Portugal central region (Serra da Estrela mountains) during the winter, in 2008. Insects

capture was held in pitfall and "Malaise" traps and specimens, subsequently, were stored individually in 70% ethanol.

All samples were morphologically identified to species level by an expert entomologist. These identifications unveiled specimens of four Diptera species: *Calliphora vicina* (13 specimens), *Calliphora vomitoria* (12 specimens), *Lucilia caesar* (8 specimens) and *Musca autumnalis* (19 specimens).

2.2 DNA extraction

DNA was extracted from 2-3 legs of each adult fly using E.Z.N.A.® Insect DNA Isolation Kit (Omega Bio-Tek, USA) following manufacturer's protocol with an overnight incubation step. To maximize final yield of DNA, 45 µL of Elution Buffer, preheated to 60 ºC – 70 ºC, was added and left to incubate for 30 - 50 minutes before centrifuging and collecting flow-through. Flow-through of the two elutions was collected in two different microtubes. Specimens' remains were retained to check their identity if necessary.

2.3 Polymerase chain reaction (PCR)

COI barcoding region was amplified using primer pair LCO1490 (5' GGTCAACAAATCATAAAGATATTGG 3') and HCO2198 (5' TAAACTTCAGGGTGACCAAAAAATCA 3') (1,2).

Each 25 µL PCR mixture contained 1X Colorless GoTaq® Flexi Reaction Buffer (Promega, USA), 100 µM of dNTPs (Fermentas, USA), 2 mM $MgCl_2$, 0.4 µM of each primer, 0.32 µg of BSA, 0.02 U GoTaq® Flexi DNA Polymerase (Promega, USA), 4-5 µL of DNA extract, and water added to complete the volume. PCR temperature cycles were carried out in a GeneAmp® PCR System 2700 thermocycler (Applied Biosystems, USA) and consisted of an initial denaturation step at 94 ºC for 1 minute, followed by 5 cycles of 94 ºC for 30 seconds, 45 ºC for 1 minute, and 72 ºC for 1 minute, and 35 cycles of 94 ºC for 1 minute, 50 ºC for 1 minute and 30 seconds, and 72 ºC for 1 minute. The last cycle

was followed by 5 minutes at 72 ºC to complete any partially synthesized strands (adapted from (1)). Amplified products were stored at 4 ºC in the original PCR mix. All PCR products checked for bands in a 1.5% agarose electrophoresis gel stained with RedSafe (iNtRON Biotechnology, Korea) through UV transillumination.

PCR products were purified with SureClean (Bioline, UK), according to manufacturer's instructions, and were stored at -20 ºC.

2.4 Sequencing

DNA was sequenced in both forward and reverse directions for all specimens using the same primers used in amplification. Sequencing reactions were performed on purified PCR products with the BigDye® Terminator v3.1 Sequencing Kit (Applied Biosystems, USA), using a GeneAmp® PCR System 2700 thermocycler. Sequencing reactions conditions consist on an inicial denaturation step at 96 ºC for 1 minute, followed by 25 cycles of 10 seconds at 96 ºC, 5 seconds at 50 ºC, and 4 minutes at 60 ºC. Then, each reaction (10 µL) was purified, transferring whole product to a clean 1.5 mL tube with 1 µL of 3 M sodium acetate, pH 4.6 and 25 µL of absolute ethanol. Mixture was then incubated in ice for 30 minutes and centrifuged at maximum speed for 25 minutes. Supernatant was discarded, 300 µL of 70% ethanol was added to the pellet, and tubes were centrifuged for another 15 minutes. This last step was repeated once, after which supernatant was discarded completely and samples air-dried away from light.

Sequencing products were then analyzed using ABI PRISM 310 Genetic Analyzer (Applied Biosystems, USA).

When this step wasn't possible to undertaken in our laboratory, samples were sent away for sequencing in a sequencing company (Macrogen Inc., Korea).

2.5 Sequence analysis

Sequence chromatograms obtained were edited and differences between forward and reverse sequences were resolved using Sequencher® v4.0.5 software (Gene Codes Corp., USA). Before analysis, all sequences were identified with GenBank BLASTn search engine (17) to confirm morphological identification. Additional COI sequence of *Hypoderma lineatum* (Viller, 1789) mitochondrial genome (accession number NC_013932) was obtained from public DNA database GenBank (18) to be used as outgroup in all analyses.

Sequences obtained in this study were aligned using ClustalX v2.0.12 (19), and BioEdit Sequence Alignment Editor v7.0.5.3 (20) was used to prepare the alignment file for posterior analyses. This file was then converted to .NEXUS format with Concatenator v1.1.0 software (21) to be used in sequence divergence and phylogenetic analyses.

Optimal model of nucleotide substitution for the data was determined using Modeltest v3.7 (22) performed in PAUP* v4.0b10 (23) according to Akaike information criterion (AIC). General time-reversible with gamma distribution shape parameter (GTR+G) model was shown as the most suitable for data analysis.

Phylogenetic analyses were carried out in PAUP* software using Maximum Parsimony (MP), Neighbor-Joining (NJ) and Maximum Likelihood (ML) methods, and in MrBayes v3.1.2 (24) for Bayesian analysis.

MP analysis was conducted using the heuristic search procedure (Tree Bisection and Reconnection algorithm, TBR) with a maxtree setting of 100 trees to find the most parsimonious trees. Bootstrap values of MP analysis (1000 replicates) were obtained under the heuristic search procedure.

A NJ tree was constructed using GTR+G model and 1000 bootstrap replicates were used to calculate support for nodes.

For ML analysis GTR+G model was also used with 1000 bootstrap replicates and 1 replicate for tree base construction.

Bayesian analysis was carried out using Monte Carlo Markov Chain method (MCMC) implemented in MrBayes (25). This Bayesian inference analysis was conducted using one cold and three hot chains, and GTR+G model, was choosed by MrModeltest v2.3 (26) as the best model for this analysis (according to AIC). During 1.500.000 generations, sampling was made every 100 generations and, to evaluate when stationary had been reached, likelihood scores from every 100 generations was plotted. From plots, it appeared that burn-in phase was completed by 30.000 generations.

To visualize tree different appearances was used TreeViewX version 0.5.1 software (27).

To study intra versus interspecific variability, uncorrected (p-distance) and corrected (Maximum Likelihood model) distances were calculated under in PAUP*, for COI fragment of 658 bp.

3. Results

A 658 bp fragment of mitochondrial COI gene was successfully amplified and sequenced for 52 different fly species.

Identification of all sequences, in the GenBank database, showed incongruences in morphological and molecular identifications of the *Musca autumnalis*. If on one hand these were previously identified as *Musca domestica*, our Blast analysis places them as *M. autumnalis*.

Aligning all sequences did not show any insertion or deletion. Data revealed 150 variable positions, of which 109 are parsimoniously informative.

3.1 Species identification

ML tree representing mitochondrial genetic differentiation of *C. vicina*, *C. vomitoria*, *L. caesar* and *M. autumnalis* species, based upon COI data, is shown in Figure 5. This tree is topologically identical to trees obtained using NJ,

MP and Bayesian methods. Phylogenetic support for individual species nodes was high (>99%) across all four methods, despite minor differences in overall topology.

Hypoderma lineatum (Diptera, Oestridae), used as outgroup, was clearly separated from Muscidae and Calliphoridae families in all analyses (Figures A1–A4, Appendix A). These two families are themselves distinct and appear monophyletic. Bootstrap values to these two families were 100% to Muscidae in all analyses and greater than 87.5% to Calliphoridae in NJ and MP analyses, despite ML analysis showed weak support (only 59.7% bootstrap). Calliphorid species were correctly assigned to sub-families Calliphorinae (*C. vicina* and *C. vomitoria*) and Luciliinae (*L. caesar*). The two species in *Calliphora* genus were grouped with high bootstrap support (>96.5% to *C. vicina* and >94.7% to *C. vomitoria*) and both species are clearly distint. Both specimens of *M. autumnalis* and *L. caesar* formed single clusters with 100% support in all analyses. Within each clade there is some variation, although this is not strongly supported by bootstrap values (<95%). Only *L. caesar3* and *L. caesar4* formed a group with bootstrap value greater than 96.8% (Figures A1–A4, Appendix A).

3.2 Intraspecific variation

Distance matrix (Table 1), based on the analysed 658 bp, revealed the percentage of nucleotide divergence values within and between among taxa. Values for intraspecific divergence with uncorrected distances (p-distance) showed a minimum of 0% for all four species and maximum reached 0.7, 0.67, 0.54 and 1.00% to *M. autumnalis*, *C. vicina*, *C. vomitoria* and *L. caesar*, respectively. Corrected distances (ML distances) revealed intraspecific divergence within four analyzed species range between 0 and 0.71% to *M. autumnalis*, 0 and 0.68% to *C. vicina*, 0 and 0.54% to *C. vomitoria* and between 0 and 1.04% to *L. caesar* (Tables A1–A2, Appendix A).

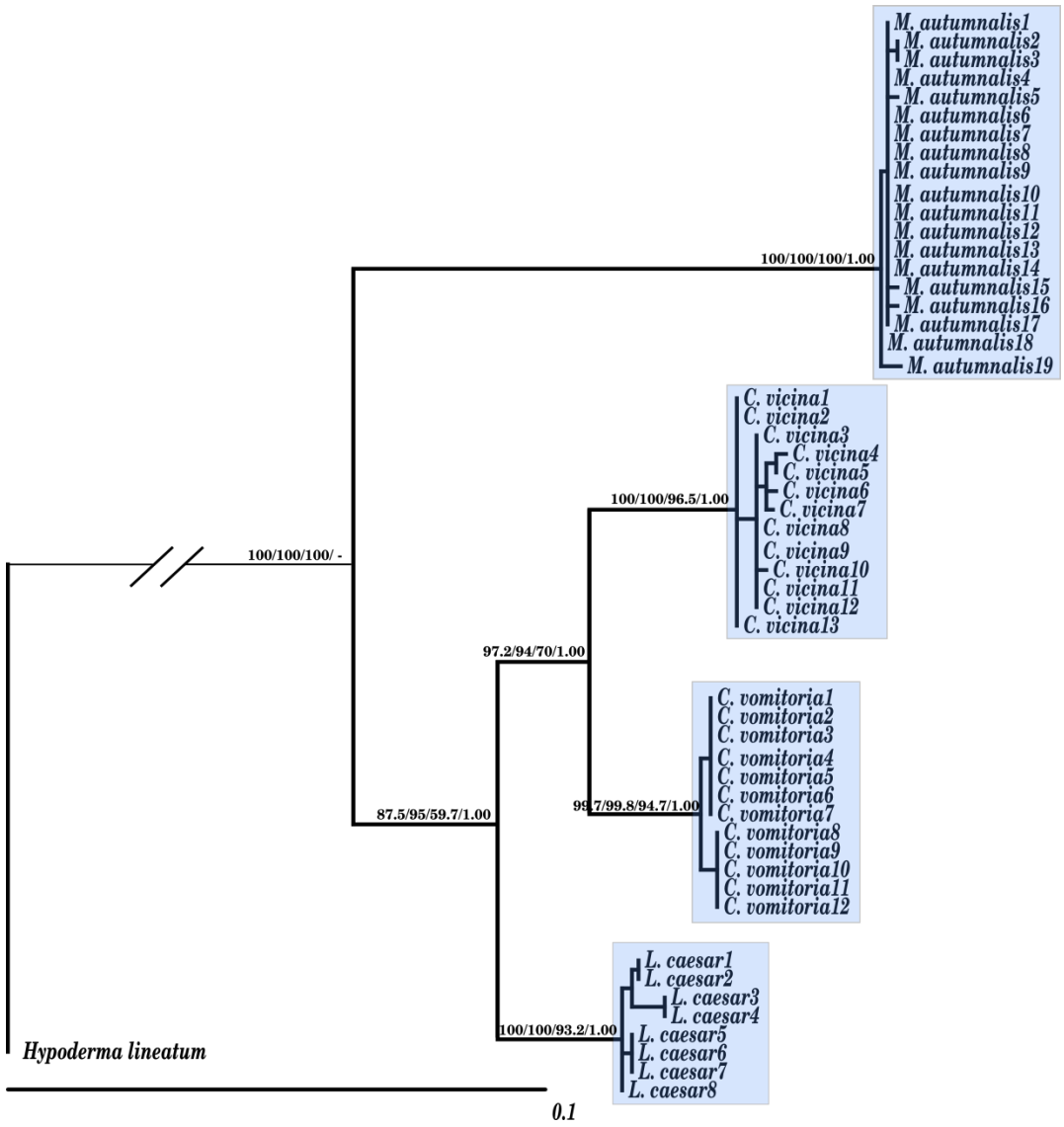

Figure 5. Maximum likelihood phylogram of 53 cytochrome *c* oxidase I (COI) sequences from four Diptera species (*Musca autumnalis*, *Calliphora vicina*, *Calliphora vomitoria* and *Lucilia caesar*) and one outgroup (*Hypoderma lineatum*). Values on tree branches correspond to Neighbor-joining/Maximum parsimony/Maximum likelihood/Bayesian inference analyses and indicate support for nodes. M = *Musca*; C = *Calliphora*; L = *Lucilia*.

3.3 Interspecific variation

Table 1 shows COI nucleotide divergence level between species groups used in analyses. Percentages of interspecific variation vary from 4.87 to 19.51% (for corrected distances) and from 3.96 to 12.01% (for p-distance).

Table 1. Percentage of divergence values within and between *Musca autumalis*, *Calliphora vicina*, *Calliphora vomitoria* and *Lucilia caesar* species at cytochrome *c* oxidase I (COI) region. Uncorrected distances (p-distances) are shown on above the diagonal and corrected distances (maximum likelihood distances) are on below the diagonal. Intraspecific divergence values are shown on the bold diagonals.

	M. autumnalis	***C. vicina***	***C. vomitoria***	***L. caesar***
M. autumnalis	**0.16**	12.01	11.05	6.19
C. vicina	19.51	**0.20**	3.96	6.19
C. vomitoria	17.87	4.87	**0.27**	5.78
L. caesar	17.41	7.73	7.10	**0.53**

In both cases, the smallest value corresponds to congeneric species, *C. vicina* and *C. vomitoria*; between *L. caesar*/*C. vicina* and *L. caesar*/*C. vomitoria* values are lower than between *M. autumnalis* and each of three other species; and highest value was found between *C. vicina* and *M. autumnalis*.

4. Discussion

The purpose of this study was to evaluate whether COI barcode provides sufficient resolution to identify different species of relevant Diptera found in forensic scenarios.

According to the DNA barcode Consortium criteria, a species identification requires monophyletic association of each species in a phylogeny (12). Here, we performed a phylogenetic analysis using four statistical methods, NJ, MP, ML and Bayesian inference which delivered each species as a monophyletic group, with strong bootstrap support (Figure 1). The high support values for each species node show the COI marker potential to be used in species discrimination, which is the fundamental premise of the DNA barcoding project. Although the COI barcode region, by itself seems do not be enough to deliver a

strong phylogenetic signal, phylogenies or resolve taxonomic associations, it seems to hold enough ability to clearly distinguish these four forensic relevant species.

The existence of a threshold value to discriminate species is another criterion used in DNA barcoding approaches. This criterion can be based on a 3% value for threshold or in a 10x or greater among versus within species nucleotide distances. In this study, intraspecific divergence within *Calliphora* species at COI region shows an average value of 0.24% (0.23% for uncorrected distances). According to 10x criterion this should correspond to a maximum sequence divergence of 2.4% (or 2.3%) as a threshold. In both cases, 2.4% and 3% thresholds, congeneric species can be distinguished, because average of interspecific divergence percentage (4.9% or 4.0%, in uncorrected distances) is greater than these two threshold values.

Additionally, it is possible to observe that the higher value of intraspecific variation correspond to *L. caesar* (Table 1). This observation confirms the apparent variation observed (with high bootstrap value) in the clade of this species (Figures A1–A4, Appendix A). Regarding the interspecific variation, lower values of divergence are observed between two congeneric species (*C. vicina* and *C. vomitoria*). Since they belong to the same genus, they are phylogenetically closest and have higher genetic similarities. Similarly, highest interspecific divergence values occur between *M. autumnalis* and the other three species. Because they belong to different families (*M. autumnalis* belongs to Muscidae; *Calliphora spp.* and *L. caesar* to Calliphoridae), these species are phylogenetically more distant.

5. Conclusions

The main aim of this study was to evaluate COI effectiveness as a marker for the correct identification of forensically relevant insects' species. Our results suggest that this COI region can be suitable for forensic relevant insects'

species identification, namely, the most common flies present. In agreement with the DNA barcoding initiative, our data shows that the use of thresholds (1,11) and monophyletic situation of species (12) allows a correct species identification.

Additionally, COI proved straightforward in amplification and sequencing. This advantage facilitates rapid generation of an unknown specimen sequence and subsequent identification. This much strengthens the use of this region as a molecular tool in forensic entomology studies and other situations featuring Diptera of applied importance.

References

1. Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. Proc R Soc Lond B 2003;270:313-321.

2. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 1994;3(5):294-299.

3. Harvey M, Dadour RI, Gaudieri S, Mansell MW, Villet MH. MtDNA-based identification of blowflies in forensic entomology: more accurate postmortem interval estimation and beyond. Forensic Sci Int 2003;136(Supplement 1):389.

4. Schroeder H, Klotzbach H, Elias S, Augustin C, Pueschel K. Use of PCR–RFLP for differentiation of calliphorid larvae (Diptera, Calliphoridae) on human corpses. Forensic Sci Int 2003;132:76-81.

5. Wells JD, Sperling FAH. DNA-based identification of forensically important Chrysomyinae (Diptera: Calliphoridae). Forensic Sci Int 2001;120:110–115.

6. Amendt J, Krettek R, Zehner R. Forensic entomology. Naturwissenschaften 2004;91:51-65.

7. Malgorn Y, Coquoz R. DNA typing for identification of some species of Calliphoridae – An interest in forensic entomology. Forensic Sci Int 1999;102:111-119.

8. Mazzanti M, Allessandrini F, Tagliabracci A, Wells JD, Campobasso CP. DNA degradation and genetic analysis of empty puparia: Genetic identification limits in forensic entomology. Forensic Sci Int 2010;195:99-102.

9. Frézal L, Leblois R. Four years of DNA barcoding: Current advances and prospects. Infect Genet Evol 2008;8:727-736.

10. Hebert PDN, Ratnasingham S, deWaard JR. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. Proc R Soc Lond B 2003;270(Supplement 1):S96-S99.

11. Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM. Identification of Birds through DNA Barcodes. PLoS Biol 2004;2(10):e312.

12. Meyer CP, Paulay G, DNA Barcoding: Error Rates Based on Comprehensive Sampling. PLoS Biol 2005;3(2):e422.

13. Whitworth TL, Dawson RD, Magalon H, Baudry E. DNA barcoding cannot reliably identify species of the blowfly genus *Protocalliphora* (Diptera: Calliphoridae). Proc R Soc B 2007;274:1731-1739.

14. Wells JD, Williams DW. Validation of a DNA-based method for identifying Chrysomyinae (Diptera: Calliphoridae) used in a death investigation. Int J Legal Med 2007;121:1-8.

15. Hajibabaei M, Singer GAC, Clare EL, Hebert PDN. Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. BMC Biol 2007;5:1-24.

16. Gusmão JB. Estudo Preliminar da Sucessão Entomológica (com Ênfase na Ordem Diptera) em Diferentes Cadáveres de Animais Selvagens na Serra da Estrela [MSc dissertation]. Lisbon: Univ. of Lisbon, 2008.

17. http://blast.ncbi.nlm.nih.gov/Blast.cgi

18. http://www.ncbi.nlm.nih.gov/nucleotide/

19. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997;25:4876-4882.

20. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 1999;41:95-98.

21. Pina-Martins F, Paulo OS. Concatenator: Sequence Data Matrices Handling Made Easy. Mol Ecol Resources 2008;8:1254-1255.

22. Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. Bioinformatics 1998;14(9):817-818.

23. Swofford DL. PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b4a — Sunderland, MA. Sinauer, 2000.

24. Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 2001;294:2310-2314.

25. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 2003;19(12):1572-1574.

26. Nylander JAA. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, 2004.

27. Page RDM. TREEVIEW: An application to display phylogenetic trees on personal computers. Comput Appl Biosci 1996;12:357-358.

Chapter 3

Forensic relevant insects' identification through GenBank and BOLD databases

CHAPTER 3 – FORENSIC RELEVANT INSECTS' IDENTIFICATION THROUGH GENBANK AND BOLD DATABASES

Abstract

The existence of entomological evidences can be of great importance to forensic cases. Indeed, this can provide relevant information to delineate the course of the investigation; therefore, the species-level identification of specimens found on corpse is extremely important. The Barcode of Life Data System (BOLD) is a new tool for management of DNA barcoding data. The identification system of BOLD is the functional unit for identification of specimens by pasting their sequence and compared this with sequence reference from known specimens, like used in others databases (e.g. GenBank from NCBI). In this way, this study arises to determine to what extent these databases are able to identify insects' species with forensic relevance. Additionally, the effectiveness of COI marker to purposes of DNA barcoding was evaluate. The results showed that GenBank allowed to identify more sequences than BOLD, and also proved the potential of COI as barcode sequence.

Keywords: forensic science; forensic entomology; database; Barcode of Life Data system; DNA barcoding; GenBank.

1. Introduction

A death body is a large food source for a range of organisms and supports a large and quickly changing fauna as it decomposes (1). Insects are generally the first organisms to colonize the corpse and they have been used as indicators

to determine postmortem interval (PMI). For forensic entomology purposes, its identification at species-level is mandatory.

The molecular genotyping methods could benefit the indispensable identification of insects' species in forensic cases. In fact, the disadvantages of the morphological identification process can be opposed by the speed and simplicity of molecular analysis, and make this the best method for forensic relevant species' identification.

In 2003, Hebert and colleagues suggest the existence of a universal sequence of DNA to identify species. This sequence, known as the barcode sequence, is the pillar for a new concept already widely spread: the DNA barcoding (2). These authors also propose a 658-bp mitochondrial genome region – the cytochrome *c* oxidase subunit I (COI) gene – as the primary barcode sequence for members of animal kingdom.

The idea of a standardized molecular identification system emerged progressively and revealed that the creation of an organization responsible by management of the DNA barcoding data would be essential. Indeed, the Consortium for the Barcode of Life (CBOL) is an international initiative that supports the development of DNA barcoding and coordinates the collection of DNA barcodes. The volume of information already existing soon after showed the necessity to build a worldwide reference database for the molecular identification of all eukaryotic species (3,4). However, that database to be a complete barcode library for the animal kingdom will have to be about 100 million records (3). In this way, CBOL initiate the construction of a new database with emphasis in DNA barcode sequences, the Barcode of Life Data System (BOLD) – www.barcodinglife.org. BOLD is a bioinformatics platform which aids the acquisition, storage, analysis and publication of DNA barcode data (3), and is a freely available resource for the DNA barcoding community. Unlike other well-known sequence depositories (e.g. GenBank from NCBI), BOLD has an interactive interface where deposited sequences can be revised and taxonomically reassigned (5).

The Identification System of this platform (BOLD-IDS) allows matching a DNA barcode sequence of an unknown specimen with an assembly of reference libraries of barcode sequences for known species. In this way, it's possible to know which species a problem-specimen belongs to. However, the recovery of species by this database could not be enough for all species discrimination. Indeed, in September 2010, the total available DNA barcode sequences were at 789 488 sequences corresponding to 75 646 species (6), a number much lower than the 100 million records previously mentioned.

In this way, this study arises to determine what extent the GenBank and BOLD databases are able to identify insects' species with forensic relevance. Additionally, we also intend to demonstrate the effectiveness of COI marker in insects' species identification.

2. Materials and Methods

2.1 Samples

The 68 samples (Table B1, Appendix B) included in this study were obtained from two previous studies (7,8). The samples were collected from vertebrate carcasses (air exposed) in Serra da Estrela Mountain (Portugal) and Oeiras (Portugal) regions between December 2007 and July 2008. The entomological material was captured with pitfall, "Malaise" and "Schoenly" traps. Then, the material was sorted, identified and stored individually in 70% ethanol.

The specimens collected were identified only at family-level because of morphological identification difficulties.

2.2 DNA extraction

DNA extraction was performed using 2 or 3 adult legs, depending on specimen size. Total genomic DNA was extracted using the E.Z.N.A.® Insect DNA Isolation Kit (Omega Bio-Tek, USA). In the first step of the procedure, samples were break down with a pestle without liquid nitrogen and, the following steps were performed according to manufacturer's protocol. However, the elution of DNA was slightly modified to ensure maximum yield, with two matrix incubations using 40 µL of Elution Buffer, preheated to 60 ºC – 70 ºC, during 30 – 50 minutes and each elution was made to a different microtube.

For purposes of DNA barcoding, some part of the specimens remains were preserved for replication of experiment if necessary.

2.3 Amplification

Initial amplification of a 658 bp 5'-end fragment of the mitochondrial COI gene was carried out using the primer pair LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') (2).

The PCR mixtures were made for a total volume of 25 µL and consisted in 1X Colorless GoTaq® Flexi Reaction Buffer (Promega, USA), 100 µM of dNTPs (Fermentas, USA), 2 mM $MgCl_2$, 0.4 µM of each primer, 0.32 µg of BSA, 0.02 U GoTaq® Flexi DNA Polymerase (Promega, USA), 4 µL of DNA, and water added to complete the final volume. Failed amplifications were repeated under the same conditions with 5 µL of genomic DNA.

PCR amplifications were performed in a GeneAmp® PCR System 2700 thermocycler (Applied Biosystems, USA), using the following conditions: 94 ºC for 1 minute, followed by 5 cycles of 94 ºC for 30 seconds, 45 ºC for 1 minute, and 72 ºC for 1 minute, 35 cycles of 94 ºC for 1 minute, 50 ºC for 1 minute and 30 seconds, and 72 ºC for 1 minute, and a final elongation for 5 minutes at 72 ºC followed by holding at 4 ºC. For some specimens amplification, the temperature

of annealing proved to be problematic and therefore optimization of the annealing temperature was obtained and used to amplify those individuals. In those cases, the PCR conditions consisted in an initial denaturation step for 1 minute at 94 ºC, 94 ºC for 1 min, 54 ºC for 1 minute, and 72 ºC for 1 minute for a total of 40 cycles, and a final elongation step for 5 minutes at 72 ºC (9).

The PCR amplicons were visualized in an agarose gel electrophoresis (1.5%), stained with RedSafe (iNtRON Biotechnology, Korea) and under UV transillumination.

2.4 Sequencing

Before sequencing, the PCR amplicons were purified with SureClean (Bioline, UK), according to manufacturer's instructions but with longer times of incubation and centrifugation, and stored at -20 ºC.

DNA sequencing was bi-directional for all specimens. The primers combination used in this step were the same used in PCR amplification. Sequencing reactions were performed using BigDye® Terminator v3.1 Sequencing Kit (Applied Biosystems, USA) according to the manufaturer's instructions. The cycle sequencing was performed in a GeneAmp® PCR System 2700 thermocycler and consist in an inicial denaturation step at 96 ºC for 1 minute, followed by 25 cycles of 10 seconds at 96 ºC, 5 seconds at 50 ºC, and 4 minutes at 60 ºC. The purification of the reaction products were made according to the following steps: transferring of reaction product to a new 1.5 mL microtube containing a solution with 1 µL of 3 M sodium acetate (pH 4.6) and 25 µL of absolute ethanol; incubate in ice during 30 minutes; centrifuge at maximum speed for 25 minutes; discard supernatant; add 300 µL of 70% ethanol to the pellet; centrifuge at maximum speed for 15 minutes; repeat the last three steps once more; discard supernatant; air-dried the samples kept in the dark.

Sequencing chromatrograms were obtained with the ABI PRISM 310 Genetic Analyzer (Applied Biosystems, USA).

2.5 Sequence analysis

Sequencing chromatograms were edited and corrected with Sequencher® v4.0.5 software (Gene Codes Corp., USA).

The specimens are molecularly identified by pasting their sequence record in both BLAST (Basic Local Alignment Search Tool) from NCBI's GenBank (10) and BOLD-IDS tool from BOLD Systems (6). In GenBank was used the nucleotide blast program for basic BLAST. The parameters used for BLAST were search in nucleotide collection database with MEGABLAST search, which is the more appropriate for comparing a query to closely related sequences. In BOLD the search was performed with BOLD-IDS tool for animal identification (that use the COI barcode) in "Species Level Barcode Records" search database and then, in "All Barcode Records on BOLD" search database when the first failed in identification.

The sequences that allowed the species-level identification were used in the next step. The alignment of Diptera sequences was carried out using the ClustalX v2.0.12 (11) and the alignment file for analysis was prepared with BioEdit Sequence Alignment Editor v7.0.5.3 (12). To avoid interferences in the analyses due to lack of some nucleotides at the beginning and end of some sequences, the sequences ends were cut. Analysis was, therefore, made with 593 bp from COI barcode fragment. To be used in sequence divergence and phylogenetic analysis the file was to be converted to .NEXUS format with Concatenator v1.1.0 program (13). The analyses was performed in PAUP* v4.0b10 (14) and in MrBayes v3.1.2 (15) software.

The optimal model of nucleotide sequence divergence for Neighbor-joining (NJ), Maximum Parsimony (MP) and Maximum Likelihood (ML) analyses, was determined using Modeltest v3.7 (16) and performed in PAUP*. According to Akaike information criterion (AIC) the General time-reversible + Proportion Invariant + Gamma distribution shape parameter (GTR+I+G) model was shown as the most suitable for the analysis. In Bayesian Inference analysis

the best model was chosen with MrModeltest v2.3 (17) and performed in MrBayes.

A NJ tree was obtained using the optimal model and the support for nodes was calculated using 1000 bootstrap replicates.

The most parsimonious tree was obtained with MP analysis using the heuristic search procedure (Tree Bisection and Reconnection algorithm, TBR) with a maxtree setting of 1000 trees. The bootstrap values were calculate with 1000 replicates and were performed under the heuristic search procedure.

For ML analysis GTR+I+G model was also used with 1000 bootstrap replicates and 10 replicates for tree base construction.

For Bayesian inference analysis, the Monte Carlo Markov Chain method (MCMC) was used in MrBayes software (18). This analysis used one cold and three heated chains with GTR+I+G model (obtain as the best model according to AIC). The sampling was made every 100 generations during 1.500.000 generations and the likelihood scores were recorded until the stationary be reached. These records shown that the burn-in phase was achieved by 30.000 generations.

The TreeViewX Version 0.5.1 software (19) was used to visualize the phylograms obtained from all analyses.

Uncorrected (p-distance) and corrected (ML) distances were calculated using the PAUP*, according to the best model previously defined, to evaluate intra and interspecific variability for the 658 bp barcode region.

3. Results

A total of 68 sequences belonging to the initial portion of mitochondrial COI gene were successfully sequenced.

The alignment of all sequences used in this study did not show any insertion or deletion.

3.1 GenBank and BOLD identifications

This study represents an effort to show the functionality and utility of species identification with a DNA barcoding marker to successful discriminate between the insects species investigated. The capacity of species identification was estimated by comparing the 68 insects sequences, analyzed for COI marker, through GenBank and BOLD databases (Table B1, Appendix B).

The Figure 6 shows the percentage of specimens identified according to each database. With GenBank database 46 of 68 samples (67.6%) was successfully identified to species-level with a maximum identity value greater than 98%. The identification was unable to 19 samples and 3 samples revealed a confused identification (the search showed two possible outcomes to the same sequence). In BOLD search, 40 sequences (58.8% of total sequences) generate a correct identification at species-level and 17 sequences (25%) identified only at genus-level with a specimen similarity value greater than 99%, for both cases. From this search has resulted 8 sequences without identification and 3 samples with confuse identification (relatively to species-level identification.

In total, 49 specimens were identified belonging to 11 diferent species: *Eudasyphora cyanella* (Meigen, 1826), *Lucilia caesar* (Linnaeus, 1758), *Pollenia rudis* (Fabricius, 1794), *Musca autumnalis* (De Geer, 1776), *Phaonia subventa* (Harris, 1780), *Phaonia tuguriorum* (Scopoli, 1763), *Helina impucta* (Fallén, 1825), *Helina evecta* (Harris, 1780), *Helina reversio* (Harris, 1780), *Hydrotaea dentipes* (Fabricius, 1805) and *Hydrotaea armipes* (Fallén, 1825).

3.2 Species identification

The ML phylogram, showing bootstrap (from NJ, MP and ML analyses) and posterior probability (obtained in Bayesian inference analysis) values, was shown in Figure 7. NJ, MP, ML and Bayesian inference performed with sequences identified to species-level showed identical tree topology (Figures B1–B4, Appendix B). *Dermestes lardarius* (Coleoptera order), used as outgroup, was

clearly separated from other specimens in all analyses. All analyses were congruent in recognizing 8 lineages on data set, almost all with high bootstrap support (in NJ, MP and ML) and posterior probability (in Bayesian inference).Only *Hydrotaea dentipes* showed lower bootstrap value (52.2%) in NJ analysis. Indeed, all species were resolved as reciprocally monophyletic groups,

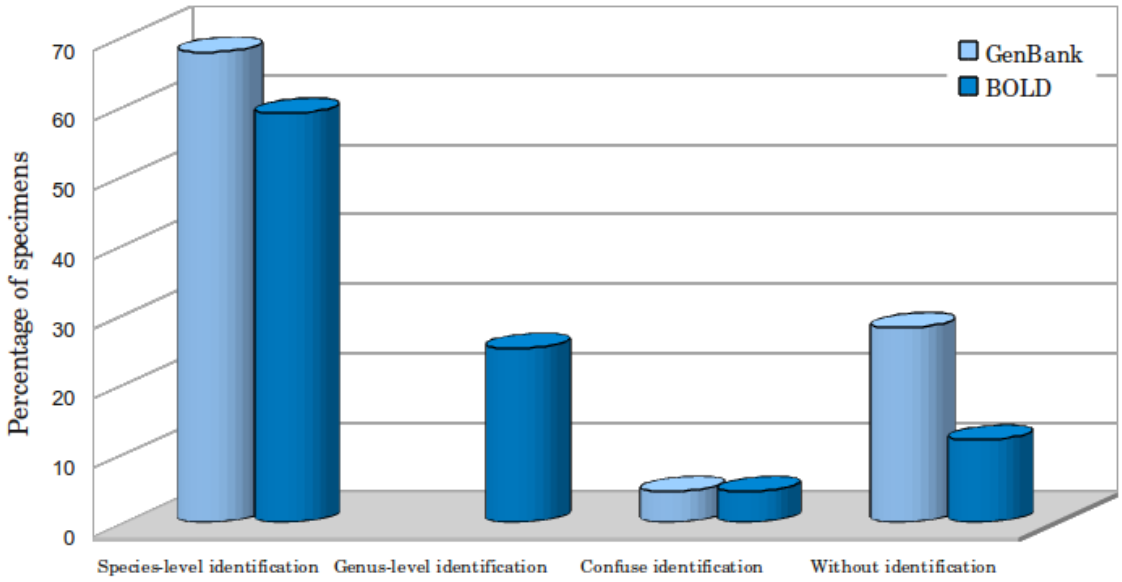

Figure 6. Percentage of specimens identified according to GenBank (dark blue bars) and BOLD (light blue bars) databases.

beside some variation can be observed within some groups. Phylogenetic analyses also indicate that *Phaonia subventa* and *Phaonia tuguriorum* never appeared associated as congeneric species. Beside this, these two congeneric species ever were shown mixed with *Helina evecta* and *Helina impucta*. *Helina evecta*/*Helina reversio* and *Helina impucta*/*Helina reversio* congeneric pairs never appears as associated at genus-level. In the other hand, *Helina impucta*/*Helina evecta* ever appear associated as congeneric species. Only NJ analysis showed association between congeneric species *Hydrotaea dentipes* and *Hydrotaea armipes*, with 100% bootstrap support. *Lucilia caesar* specimen was showed alone in all analyses.

Table 2 compares the percentage of intraspecific and interspecific nucleotide divergences between congeneric species. Comparing these values it's

possible to observe that, all intraspecific values are lower than 3% and the interspecific percentages are much higher than this value. In the other hand, all genera present an interspecific divergence percentage greater than its 10x intraspecific divergence percentage.

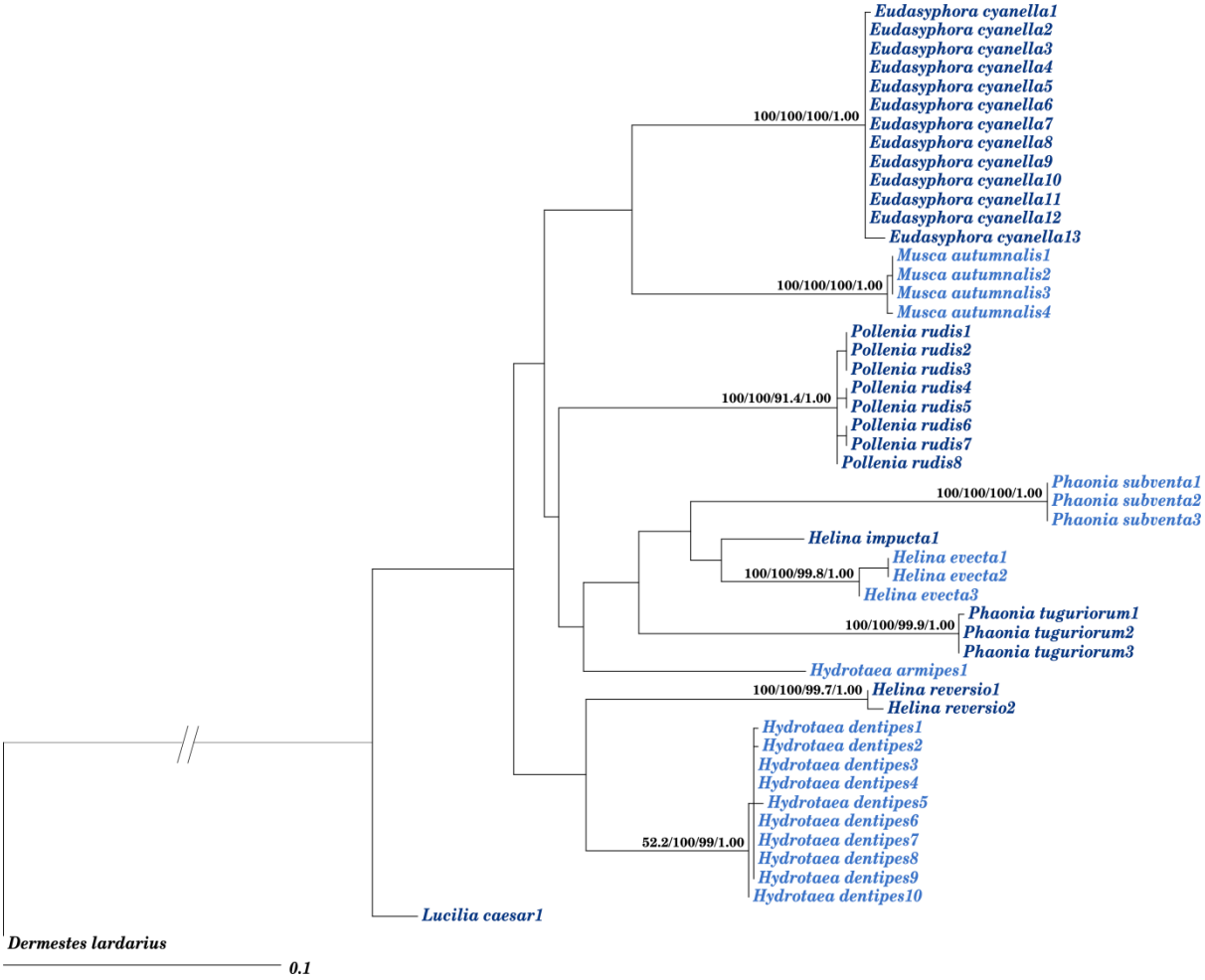

Figure 7. Maximum likelihood phylogram of 69 cytochrome *c* oxidase I (COI) sequences from ten Diptera species and one outgroup (*Dermestes lardarius*). Values on tree branches correspond to Neighbor-joining/Maximum parsimony/Maximum likelihood/Bayesian inference and indicate support for nodes.

Table 2. Summary of intra and interspecific percentages of nucleotide divergences at cytochrome *c* oxidase I (p-distances and ML distances) of *Phaonia*, *Helina* and *Hydrotaea* genera.

	p-distances		ML distances	
Genus	**Intraspecific divergence**	**Interspecific divergence**	**Intraspecific divergence**	**Interspecific divergence**
Phaonia	***0.05***	***14.55***	***0.05***	***16.81***
Helina	***0.63***	***9.96***	***0.64***	***11.16***
Hydrotaea	***0.16***	***8.07***	***0.16***	***8.83***

4. Discussion

The comparison between the two molecular databases, GenBank and BOLD, reveals that GenBank database can identify more query sequences than BOLD database. This can be due to the fact that GenBank presents a most comprehensive database than BOLD (this is a more recent and specific database). Other fact can be associated with the BLAST search tools. These databases use different algorithms to calculate the similarity between reference and query sequences, and this can generate discrepancies in identification. In GenBank search, the 98% was used as limit in species identification because was observed that values below this delivery the query sequences to a different species than the species showed with values greater than 98%. In BOLD search, this value was 99%. According to this database species level match could not be made with values lower than 99%, returning only the information which is the nearest neighbor species.

Comparing the performance of these four tree-building methods it is possible considered that all give similar results, recovering each species as a monophyletic group. Moreover, almost all bootstrap and posterior probability values were high showing the potential of this genetic marker to be used as a trustworthy marker in species identification.

However, some handicaps were observed in phylogenetic analysis. The non-association revealed between some congeneric species questions the power of this marker. Unfortunately, insufficient sequences of some species were available for a more detailed analysis, and the lack of some information in beginning and end of sequences may have interfered giving non-realistic results (considering that the species taxonomic level are well defined). The outgroup choice may also have interfered with the phylogenetic structure we would expect with this dataset. This may be a too distant outgroup to give rise to a tree more clearly defined.

The mean of intraspecific and interspecific variation values were calculated only when two or more congeneric species exist. Keeping in attention the threshold values given for species discrimination, 3% (2) and 10x intraspecific divergence mean for each genus (20), the results showed that was possible distinguish the two species of *Phaonia* (*Phaonia subventa* and *Phaonia tuguriorum*), the two species of *Hydrotaea* (*Hydrotaea dentipes* and *Hydrotaea armipes*), and the three species of *Helina* (*Helina impucta*, *Helina evecta* and *Helina reversion*). Indeed, 0.05%, 0.64% and 0.16% of intraspecific variations means for *Phaonia*, *Helina* and *Hydrotaea* are lower than 3% threshold. In the other hand, reveal a threshold value of 0.5%, 6.4% and 1.6%, respectively (values calculated by 10x rule) and, in all cases, these values were lower than means of interspecific variation (16.81% for *Phaonia*, 11.16% for *Helina* and 8.83% for *Hydrotaea*).

5. Conclusions

The greatest approach to identify an unidentified sequence is to notice if that sequence already exists in a public database. The identification of Diptera's species with forensic relevance showed to be of extremely importance for the investigation progress. As main aim, this study arises to determine what extent the GenBank and BOLD databases are able to identify these species. It was

possible to determine that these two databases allow identify a good percentage of species with forensic interest. However, any effort that contributes to a better understanding of biodiversity (in particular, with forensic interest; in general, for the biodiversity quantification) is of utmost importance, and the implementation of a new database comprehensive to this part of biological diversity, it's a good step in direction to this knowledge. The establishment of a standard protocol may contribute to faster growth of this database. Consequently, here we also tested the effectiveness of COI barcode to be used in a standard protocol. The results support the potential of this genetic marker. However, more comprehensive studies should be developed, with more samples and others genetic markers, to overcome some difficulties encountered in this study.

References

1. Özdemir S, Sert O. Determination of Coleoptera fauna on carcasses in Ankara province, Turkey. Forensic Sci Int 2003;183:24–32.
2. Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. Proc R Soc Lond B 2003;270:313-321.
3. Ratnasingham S, Hebert PDN. BOLD: The Barcode of Life Data System (www.barcodinglife.org). Mol Ecol Notes 2007;7(3):355-364.
4. Valentini A, Pompanon F, Taberlet P. DNA barcoding for ecologists. Trends Ecol Evol 2009;24(2):110-117.
5. Frézal L, Leblois R. Four years of DNA barcoding: Current advances and prospects. Infect Genet Evol 2008;8:727-736.
6. www.boldsystems.org
7. Gusmão JB. Estudo Preliminar da Sucessão Entomológica (com Ênfase na Ordem Diptera) em Diferentes Cadáveres de Animais Selvagens na Serra da Estrela [dissertation]. Lisbon: Univ. of Lisbon, 2008.

8. Marques AMA. Entomologia Forense: Análise da Entomofauna em Cadáveres de *Sus scrofa* (Linnaeus), na Região de Oeiras, Portugal [dissertation]. Lisbon: Univ. of Lisbon, 2008.

9. Hajibabaei M, Singer GAC, Clare EL, Hebert PDN. Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. BMC Biol 2007;5:1-24.

10. http://blast.ncbi.nlm.nih.gov

11. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997;25:4876-4882.

12. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 1999;41:95-98.

13. Pina-Martins F, Paulo OS. Concatenator: Sequence Data Matrices Handling Made Easy. Mol Ecol Resources 2008;8:1254-1255.

14. Swofford DL. PAUP*: Phylogenetic analysis using parsimony (* and other methods), Version 4.0b4a — Sunderland, MA. Sinauer, 2000.

15. Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 2001;294:2310-2314.

16. Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. Bioinformatics 1998;14(9):817-818.

17. Nylander JAA. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, 2004.

18. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 2003;19(12):1572-1574.

19. Page RDM. TREEVIEW: An application to display phylogenetic trees on personal computers. Comput Appl Biosci 1996;12:357-358.

20. Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM. Identification of Birds through DNA Barcodes. PLoS Biol 2004;2(10):e312.

Chapter 4

Final Considerations

FINAL CONSIDERATIONS

This study was the first molecular approach to assessing the potential of DNA barcoding, especially of COI marker for its inclusion in a database of species of forensic interest. In addition, a database of these adds knowledge of biodiversity that can be used in other situations of ecological and conservationist context. Indeed, Portugal is a country with a very particular geoclimatic condition, and the survey of their biodiversity is extremely important because it can reveal some unknown endemic species, and thus contribute to the global understanding of biological diversity.

In this study, morphological identification was overpass by this molecular approach in that morphological identification revealed a weakness in identification of some species. The weakness of the morphological methodologies refers mainly to the difficulty of observation of some morphological characters of identification which can lead to an incorrect identification. Moreover, this weakness reinforces the importance of molecular identification.

The successful amplification and sequencing of COI marker showed its potential to be used in a standard protocol that quickly allows obtain the sequences and subsequent identification of species. The importance of using a well-supported protocol to be used as standard protocol in forensic investigation services will facilitate the course of the investigation both in the context of forensic medicine, whether in the context of attacks on wildlife destruction.

Appendix A

(regarding to Chapter 2)

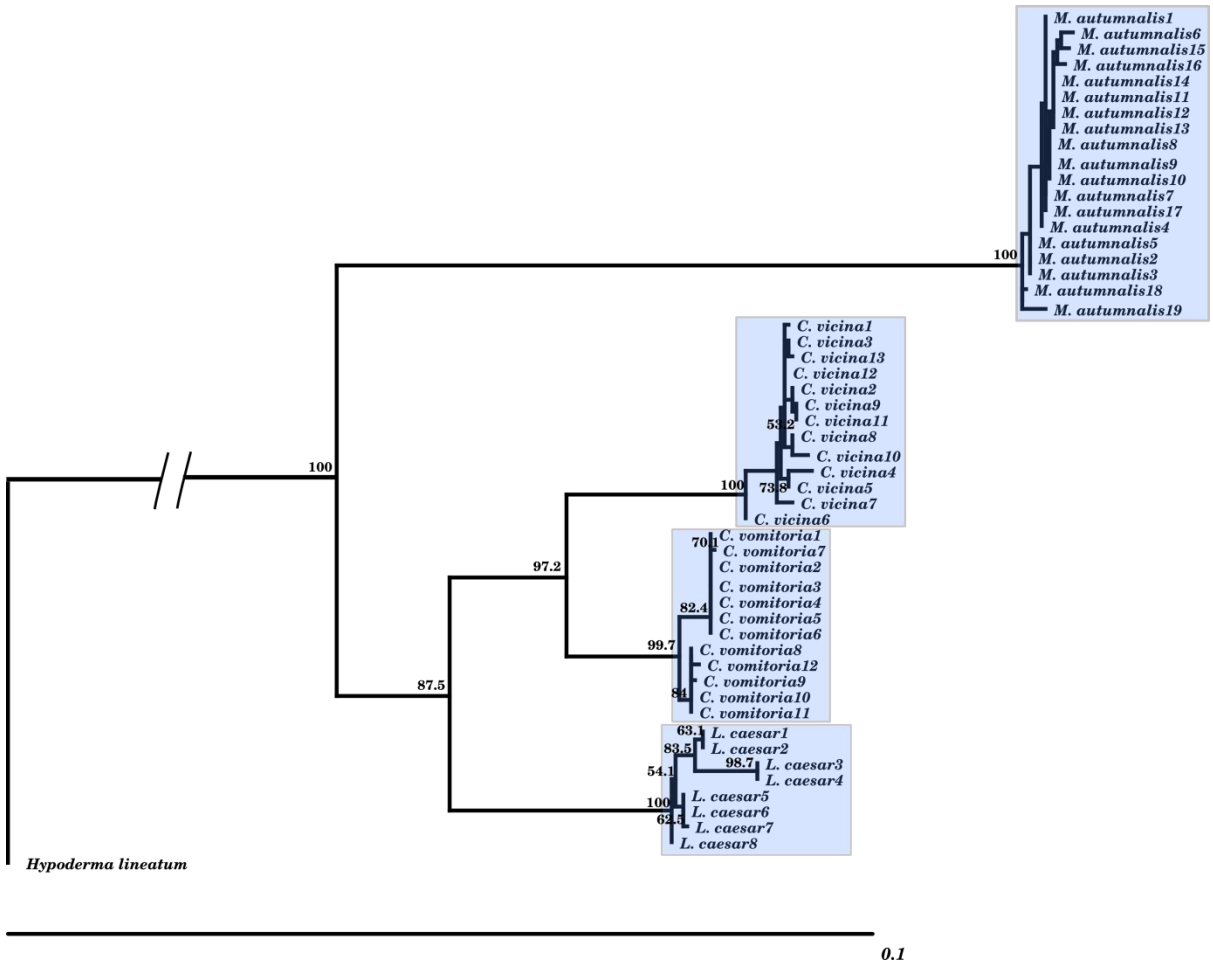

Figure A1. Neighbor-joining phylogram of 53 cytochrome *c* oxidase I (COI) sequences from four Diptera species (*Musca autumnalis*, *Calliphora vicina*, *Calliphora vomitoria* and *Lucilia caesar*) and one outgroup (*Hypoderma lineatum*). Bootstrap values indicate support for nodes among 1000 bootstrap replicates. M = *Musca*; C = *Calliphora*; L = *Lucilia*.

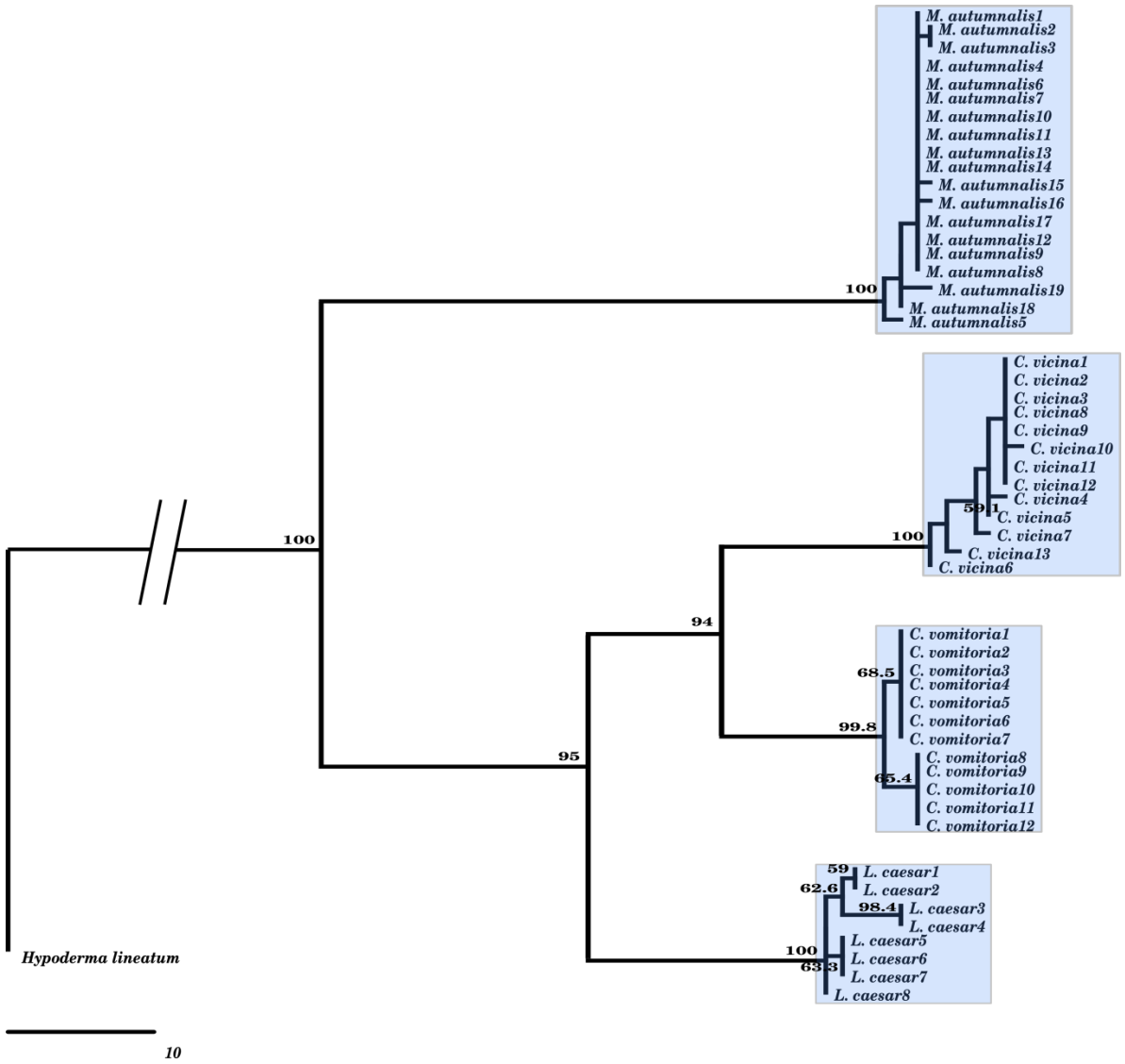

Figure A2. Maximum parsimony phylogram of heuristic search procedure (Tree Bisection and Reconnection algorithm, TBR) for 53 cytochrome *c* oxidase I (COI) sequences from four Diptera species (*Musca autumnalis*, *Calliphora vicina*, *Calliphora vomitoria* and *Lucilia caesar*) and one outgroup (*Hypoderma lineatum*). Bootstrap values indicate support for nodes among 1000 bootstrap replicates (heuristic search procedure). M = *Musca*; C = *Calliphora*; L = *Lucilia*.

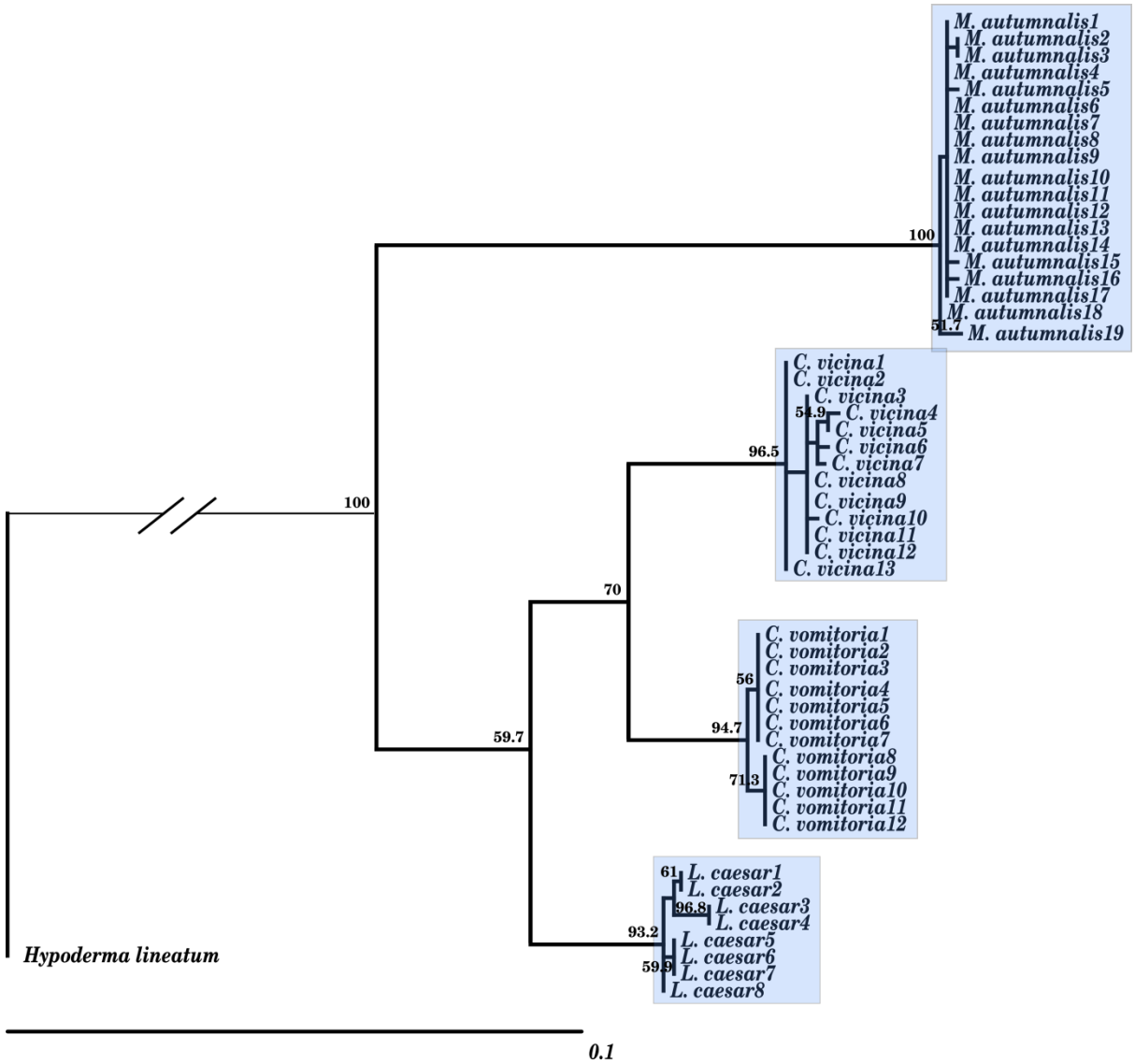

Figure A3. Maximum likelihood phylogram of 53 cytochrome *c* oxidase I (COI) sequences from four Diptera species (*Musca autumnalis*, *Calliphora vicina*, *Calliphora vomitoria* and *Lucilia caesar*) and one outgroup (*Hypoderma lineatum*). Bootstrap values indicate support for nodes among 1000 bootstrap replicates. M = *Musca*; C = *Calliphora*; L = *Lucilia*.

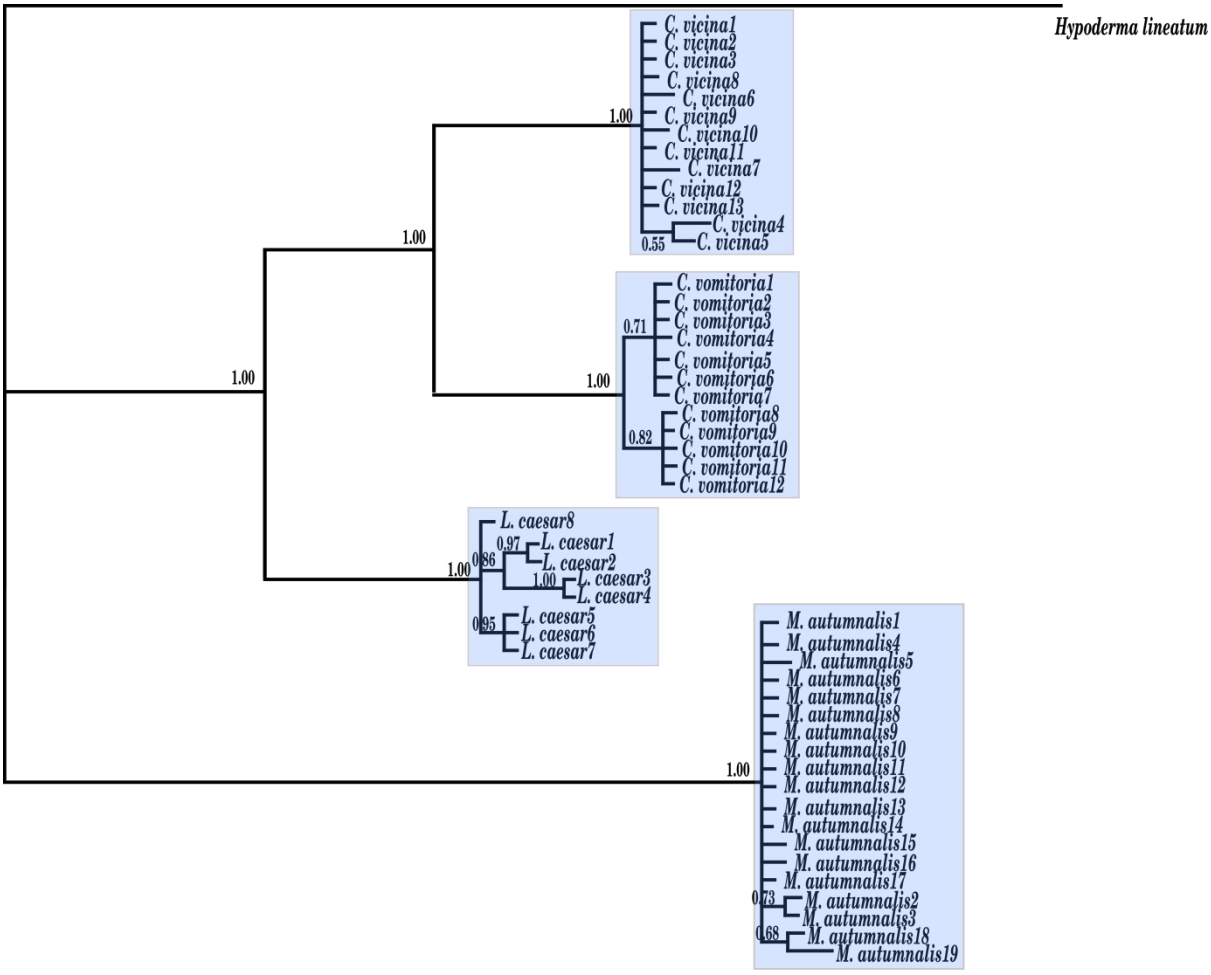

Figure A4. Bayesian phylogeny of 53 cytochrome *c* oxidase I (COI) sequences from four Diptera species (*Musca autumnalis*, *Calliphora vicina*, *Calliphora vomitoria* and *Lucilia caesar*) and one outgroup (*Hypoderma lineatum*). Values on tree branches indicate posterior probability for nodes. M = *Musca*; C = *Calliphora*; L = *Lucilia*.

Table A1. Percentage of nucleotide sequence divergence (p-distances) at cytochrome c oxidase I (COI) region for *Musca autumalis* , *Calliphora vicina* , *Calliphora vomitoria* and *Lucilia caesar* species. M = *Musca* ; C = *Calliphora* ; L = *Lucilia*; *H* = *Hypoderma* .

	M. autumnalis1	*M. autumnalis2*	*M. autumnalis4*	*M. autumnalis5*	*M. autumnalis6*	*M. autumnalis7*	*M. autumnalis18*	*M. autumnalis8*	*M. autumnalis9*	*M. autumnalis10*	*M. autumnalis11*	*M. autumnalis12*	*M. autumnalis13*	*M. autumnalis14*	*M. autumnalis15*	*M. autumnalis16*	*M. autumnalis17*	*M. autumnalis3*	*M. autumnalis19*	*C. vicina1*	*C. vicina2*	*C. vicina3*	*C. vicina4*	*C. vicina8*	*C. vicina5*	*C. vicina6*	*C. vicina9*	*C. vicina10*	*C. vicina11*	*C. vicina7*	*C. vicina12*	*C. vicina13*	*C. vomitoria1*	*C. vomitoria8*	*C. vomitoria2*	*C. vomitoria3*	*C. vomitoria9*	*C. vomitoria4*	*C. vomitoria5*	*C. vomitoria6*	*C. vomitoria10*	*C. vomitoria11*	*C. vomitoria7*	*C. vomitoria12*	*L. caesar1*	*L. caesar3*	*L. caesar5*	*L. caesar6*	*L. caesar2*	*L. caesar8*	*L. caesar4*	*L. caesar7*	*H. lineatum*
M. autumnalis1		0,2	0,0	0,2	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,2	0,0	0,2	0,5	12,2	12,1	12,0	12,5	11,9	12,1	11,3	11,9	12,0	11,9	11,9	11,9	11,8	10,9	10,5	11,0	11,0	10,7	11,0	11,0	11,0	10,7	11,0	11,0	10,7	11,2	11,7	11,0	11,0	11,2	10,8	11,7	11,0	15,6
M. autumnalis2			0,2	0,3	0,2	0,2	0,3	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,3	0,3	0,2	0,0	0,7	12,1	12,0	11,8	12,3	11,8	11,9	11,1	11,7	11,9	11,7	11,7	11,7	11,7	10,7	10,3	10,8	10,8	10,5	10,9	10,9	10,9	10,5	10,9	10,9	10,5	11,0	11,5	10,8	10,8	11,0	10,6	11,5	10,8	15,6
M. autumnalis4				0,2	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,2	0,0	0,2	0,5	12,2	12,1	12,0	12,5	11,9	12,1	11,2	11,9	12,0	11,9	11,8	11,9	11,8	10,9	10,5	11,0	11,0	10,6	11,0	11,0	11,0	10,7	11,0	11,0	10,7	11,2	11,7	10,9	10,9	11,1	10,8	11,6	11,0	15,6
M. autumnalis5					0,2	0,2	0,3	0,2	0,2	0,2	0,2	0,2	0,2	0,3	0,3	0,2	0,3	0,7	12,0	12,0	11,8	12,3	11,7	11,9	11,1	11,7	11,9	11,7	11,7	11,7	11,6	10,9	10,5	11,0	11,0	10,6	11,0	11,0	11,0	10,7	11,0	11,0	10,7	11,0	11,5	10,8	10,8	11,0	10,6	11,5	10,8	15,6
M. autumnalis6						0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,2	0,0	0,2	0,5	12,2	12,3	12,3	12,5	12,1	12,4	11,6	12,3	12,5	12,3	12,1	12,3	12,3	11,4	11,0	11,6	11,6	11,2	11,6	11,6	11,6	11,2	11,2	11,6	11,2	11,6	12,1	11,3	11,3	11,5	11,2	12,1	11,4	16,2
M. autumnalis7							0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,2	0,0	0,2	0,5	12,2	12,1	12,0	12,5	12,0	12,1	11,3	12,0	12,1	12,0	12,0	12,0	11,9	11,0	10,6	11,1	11,1	10,8	11,1	11,1	11,1	10,8	11,1	11,1	10,8	11,1	11,6	10,9	10,9	11,1	10,7	11,6	10,9	15,8
M. autumnalis18								0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,3	0,3	0,2	0,3	0,3	12,1	12,0	11,8	12,3	11,8	11,9	11,1	11,7	11,9	11,7	11,7	11,7	11,7	10,7	10,3	10,9	10,9	10,5	10,9	10,9	10,9	10,5	10,9	10,9	10,5	11,1	11,6	10,8	10,8	11,0	10,7	11,5	10,9	15,5
M. autumnalis8									0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,2	0,0	0,2	0,5	12,2	12,1	12,0	12,5	11,9	12,1	11,3	12,2	12,3	12,2	12,1	11,9	11,9	11,2	10,8	11,3	11,3	10,9	11,3	11,3	11,3	11,0	11,0	11,3	10,9	11,1	11,6	10,9	10,9	11,1	10,7	11,6	10,9	15,6
M. autumnalis9										0,0	0,0	0,0	0,0	0,0	0,2	0,2	0,0	0,2	0,5	12,2	12,1	12,0	12,5	11,9	12,1	11,3	12,2	12,3	12,2	12,1	11,9	11,9	11,2	10,8	11,3	11,3	11,3	10,9	11,3	11,3	11,3	11,0	11,0	11,3	10,9	11,1	11,6	10,9	10,9	11,1	10,7	11,6	10,9	15,6
M. autumnalis10											0,0	0,0	0,0	0,0	0,2	0,2	0,0	0,2	0,5	12,2	12,1	12,0	12,5	11,9	12,1	11,2	12,2	12,3	12,2	12,1	11,9	11,8	11,0	10,6	11,1	11,1	10,8	11,2	11,2	11,2	10,8	11,0	11,1	10,8	11,1	11,6	10,9	10,9	11,1	10,7	11,6	10,9	15,5
M. autumnalis11												0,0	0,0	0,0	0,2	0,2	0,0	0,2	0,5	12,2	12,1	12,0	12,5	11,9	12,0	11,2	12,2	12,3	12,1	12,1	11,9	12,0	11,1	10,8	11,4	11,4	11,0	11,3	11,3	11,3	10,9	11,0	11,4	11,0	11,2	11,7	11,0	11,0	11,0	10,7	11,7	11,0	15,7
M. autumnalis12													0,0	0,0	0,2	0,2	0,0	0,2	0,5	12,2	12,1	12,0	12,5	11,9	12,0	11,2	12,2	12,3	12,1	12,1	11,9	12,0	11,1	10,8	11,4	11,4	11,0	11,3	11,3	11,3	10,9	11,0	11,4	11,0	11,2	11,7	11,0	11,0	11,0	10,7	11,7	11,0	15,7
M. autumnalis13														0,0	0,2	0,2	0,0	0,2	0,5	12,2	12,1	12,0	12,5	11,9	12,0	11,2	12,2	12,3	12,1	12,1	11,9	12,0	11,1	10,8	11,4	11,4	11,0	11,3	11,3	11,3	10,9	11,0	11,4	11,0	11,2	11,7	11,0	11,0	11,0	10,7	11,7	11,0	15,7
M. autumnalis14															0,2	0,2	0,0	0,2	0,5	12,2	12,1	12,0	12,5	12,0	12,1	11,3	12,2	12,4	12,2	12,1	11,9	12,0	11,2	10,8	11,4	11,4	11,1	11,3	11,3	11,3	11,0	11,0	11,6	11,3	11,1	11,6	10,9	10,9	10,9	10,6	11,7	11,0	15,4
M. autumnalis15																0,3	0,2	0,3	0,7	12,4	12,3	12,2	12,7	12,2	12,3	11,5	12,4	12,5	12,3	12,3	12,1	12,2	11,4	11,0	11,6	11,6	11,2	11,5	11,5	11,5	11,1	11,2	11,7	11,4	11,3	11,7	11,1	11,1	11,1	10,8	11,9	11,2	15,5
M. autumnalis16																	0,2	0,3	0,7	12,4	12,3	12,2	12,7	12,2	12,3	11,5	12,4	12,5	12,3	12,3	12,1	12,2	11,4	11,0	11,6	11,6	11,2	11,5	11,5	11,5	11,1	11,2	11,7	11,4	11,1	11,6	11,1	11,1	10,9	10,8	11,7	11,2	15,4
M. autumnalis17																		0,2	0,5	12,2	12,1	12,0	12,5	11,9	12,1	11,3	12,2	12,3	12,1	12,1	11,9	11,8	11,2	10,8	11,3	11,3	10,9	11,3	11,3	11,3	10,9	11,0	11,3	11,0	10,8	11,3	10,7	10,7	10,9	10,6	11,5	10,9	15,0
M. autumnalis3																			0,7	12,0	12,0	11,8	12,3	11,7	11,9	11,1	12,0	12,2	12,0	11,9	11,7	11,9	11,0	10,6	11,1	11,1	10,7	11,1	11,1	11,1	10,8	10,8	11,3	10,9	10,9	11,4	10,6	10,6	10,7	10,4	11,4	10,7	15,3
M. autumnalis19																				12,2	12,2	12,0	12,7	12,1	12,2	11,4	11,9	11,9	11,9	11,9	11,9	11,9	11,1	10,7	11,0	11,0	10,7	11,0	11,0	11,0	10,7	11,0	11,0	10,7	11,2	11,7	10,8	10,8	11,1	10,8	11,7	10,8	15,4
C. vicina1																					0,0	0,0	0,4	0,0	0,2	0,4	0,0	0,2	0,0	0,0	0,0	0,2	3,8	4,0	3,7	3,7	3,9	3,7	3,7	3,7	3,9	3,9	3,7	3,9	6,1	6,6	5,8	5,8	6,0	5,6	6,6	5,8	13,0
C. vicina2																						0,0	0,4	0,0	0,2	0,4	0,0	0,2	0,0	0,2	0,0	0,0	3,9	4,0	3,8	3,8	4,0	3,8	3,8	3,8	4,0	4,0	3,8	4,0	6,1	6,6	5,9	5,9	6,0	5,7	6,6	5,9	12,8
C. vicina3																							0,4	0,0	0,2	0,4	0,0	0,2	0,0	0,3	0,0	0,0	3,8	4,0	3,8	3,8	3,9	3,8	3,8	3,8	3,9	4,0	3,8	3,9	6,0	6,5	5,8	5,8	6,0	5,6	6,5	5,8	12,8
C. vicina4																								0,4	0,2	0,6	0,4	0,4	0,4	0,4	0,4	0,4	3,9	4,1	3,9	3,9	4,1	3,9	3,9	3,9	4,1	4,1	3,9	4,1	6,4	6,9	5,9	5,9	6,3	5,9	6,9	5,9	13,3
C. vicina8																									0,2	0,4	0,0	0,0	0,0	0,2	0,0	0,0	4,0	4,2	4,0	4,0	4,2	4,0	4,0	4,0	4,2	4,2	4,0	4,2	5,9	6,3	5,5	5,5	5,9	5,5	6,3	5,5	12,7
C. vicina5																										0,4	0,2	0,2	0,2	0,4	0,2	0,2	3,8	4,0	3,8	3,8	4,0	3,8	3,8	3,8	4,0	4,1	3,8	4,0	6,1	6,7	5,7	5,7	6,1	5,7	6,6	5,7	12,9
C. vicina6																											0,4	0,4	0,4	0,4	0,4	0,4	3,5	3,6	3,4	3,4	3,6	3,4	3,4	3,4	3,6	3,7	3,5	3,6	6,4	6,9	6,0	6,0	6,3	6,0	6,9	6,0	12,9
C. vicina9																												0,2	0,0	0,3	0,0	0,3	4,1	4,3	4,1	4,1	4,2	4,1	4,1	4,1	4,2	4,0	4,1	4,2	6,4	7,0	6,2	6,2	6,4	6,0	6,9	6,2	13,0
C. vicina10																													0,2	0,5	0,2	0,5	4,3	4,4	4,2	4,2	4,4	4,2	4,2	4,2	4,4	4,2	4,2	4,4	6,6	7,1	6,1	6,0	6,6	6,2	7,1	6,1	12,9
C. vicina11																														0,3	0,0	0,3	4,1	4,3	4,0	4,0	4,2	4,0	4,0	4,0	4,2	4,0	4,1	4,2	6,4	6,9	6,2	6,2	6,4	6,0	6,9	6,2	13,0
C. vicina7																															0,3	0,7	4,1	4,3	4,0	4,0	4,2	4,0	4,0	4,0	4,2	3,8	4,0	4,2	6,4	6,9	6,2	6,2	6,4	6,0	6,9	6,2	12,8
C. vicina12																																0,0	3,8	4,0	3,7	3,7	3,9	3,7	3,7	3,7	3,9	4,0	3,7	3,9	6,1	6,6	5,9	5,9	6,1	5,7	6,6	5,9	12,7
C. vicina13																																	4,4	4,3	4,0	4,0	3,9	3,9	3,9	4,2	4,0	4,0	4,1	4,0	6,1	6,5	6,0	6,1	6,1	5,8	6,9	6,2	12,8
C. vomitoria1																																																					12,8
C. vomitoria8																																																					13,0
C. vomitoria2																																																				12,6
C. vomitoria3																																																				12,6
C. vomitoria9																																																				12,7
C. vomitoria4																																																				12,6
C. vomitoria5																																																				12,6
C. vomitoria6																																																				12,6
C. vomitoria10																																																				12,8
C. vomitoria11																																																				13,0
C. vomitoria7																																																				12,8
C. vomitoria12																																																				12,9
L. caesar1																																																				12,6
L. caesar3																																																				13,1
L. caesar5																																																				12,4
L. caesar6																																																				12,4
L. caesar2																																																				12,4
L. caesar8																																																				12,4
L. caesar4																																																				13,2
L. caesar7																																																				12,6
H. lineatum																																																				

Table A2. Percentage of nucleotide sequence divergence (ML distances) at cytochrome c oxidase I (COI) region for *Musca autumalis*, *Calliphora vicina*, *Calliphora vomitoria* and *Lucilia caesar* species. M = *Musca*; C = *Calliphora*; L = *Lucilia*; H = *Hypoderma*.

	M. autumnalis1	*M. autumnalis2*	*M. autumnalis4*	*M. autumnalis5*	*M. autumnalis6*	*M. autumnalis7*	*M. autumnalis18*	*M. autumnalis8*	*M. autumnalis9*	*M. autumnalis10*	*M. autumnalis11*	*M. autumnalis12*	*M. autumnalis13*	*M. autumnalis14*	*M. autumnalis15*	*M. autumnalis16*	*M. autumnalis17*	*M. autumnalis3*	*M. autumnalis19*	*C. vicina1*	*C. vicina2*	*C. vicina3*	*C. vicina4*	*C. vicina8*	*C. vicina5*	*C. vicina6*	*C. vicina9*	*C. vicina10*	*C. vicina11*	*C. vicina7*	*C. vicina12*	*C. vicina13*	*C. vomitoria1*	*C. vomitoria8*	*C. vomitoria2*	*C. vomitoria3*	*C. vomitoria9*	*C. vomitoria4*	*C. vomitoria5*	*C. vomitoria6*	*C. vomitoria10*	*C. vomitoria11*	*C. vomitoria7*	*C. vomitoria12*	*L. caesar1*	*L. caesar3*	*L. caesar5*	*L. caesar6*	*L. caesar2*	*L. caesar8*	*L. caesar4*	*L. caesar7*	*H. lineatum*
M. autumnalis1																																																					
M. autumnalis2	0,2																																																					
M. autumnalis4	0,0	0,2																																																		
M. autumnalis5	0,2	0,3	0,2																																																	
M. autumnalis6	0,0	0,2	0,0	0,2																																																
M. autumnalis7	0,0	0,2	0,0	0,2	0,0																																															
M. autumnalis18	0,2	0,4	0,2	0,4	0,2	0,2																																														
M. autumnalis8	0,0	0,2	0,0	0,2	0,0	0,0	0,2																																													
M. autumnalis9	0,0	0,2	0,0	0,2	0,0	0,0	0,2	0,0																																											
M. autumnalis10	0,0	0,2	0,0	0,2	0,0	0,0	0,2	0,0	0,0																																									
M. autumnalis11	0,0	0,2	0,0	0,2	0,0	0,0	0,2	0,0	0,0	0,0																																								
M. autumnalis12	0,0	0,2	0,0	0,2	0,0	0,0	0,2	0,0	0,0	0,0	0,0																																							
M. autumnalis13	0,0	0,2	0,0	0,2	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0																																							
M. autumnalis14	0,0	0,2	0,0	0,2	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0																																						
M. autumnalis15	0,2	0,3	0,2	0,3	0,2	0,2	0,3	0,2	0,2	0,2	0,2	0,2	0,2	0,2																																						
M. autumnalis16	0,2	0,3	0,2	0,3	0,2	0,2	0,3	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,3																																						
M. autumnalis17	0,0	0,2	0,0	0,2	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,2																																				
M. autumnalis3	0,2	0,0	0,2	0,3	0,2	0,2	0,4	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,3	0,3	0,2																																				
M. autumnalis19	0,5	0,7	0,5	0,7	0,6	0,5	0,4	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,7	0,7	0,5	0,7																																				
C. vicina1	20,0	19,6	20,0	19,6	20,0	20,0	19,6	20,0	20,0	20,0	20,0	20,0	20,0	20,0	20,4	20,4	20,0	19,6	20,3																																				
C. vicina2	19,8	19,3	19,8	19,3	20,0	19,8	19,3	19,8	19,8	19,8	19,8	19,8	19,8	19,8	20,1	20,1	19,8	19,3	20,0	0,0																																				
C. vicina3	19,4	19,0	19,4	19,0	20,0	19,4	19,0	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,7	19,7	19,4	19,0	19,6	0,0	0,0																																				
C. vicina4	20,8	20,3	20,8	20,3	20,9	20,8	20,3	20,8	20,8	20,8	20,8	20,8	20,8	20,8	21,2	21,2	20,8	20,3	21,3	0,4	0,4	0,4																																				
C. vicina8	19,4	18,9	19,4	18,9	19,7	19,4	18,9	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,8	19,8	19,4	18,9	19,9	0,0	0,0	0,0	0,4																																				
C. vicina5	19,6	19,2	19,6	19,2	20,4	19,7	19,2	19,6	19,6	19,6	19,6	19,6	19,6	19,6	20,0	20,0	19,6	19,2	20,1	0,2	0,2	0,2	0,2	0,2																																				
C. vicina6	17,9	17,4	17,9	17,4	18,7	18,1	17,4	17,9	17,9	17,9	17,9	17,9	17,9	17,9	18,2	18,2	17,9	17,4	18,3	0,4	0,4	0,4	0,6	0,4	0,4																																				
C. vicina9	19,1	18,7	19,1	18,7	20,0	19,4	18,8	19,8	19,8	19,8	19,8	19,8	19,8	19,8	20,1	20,1	19,8	19,4	19,3	0,0	0,0	0,0	0,4	0,0	0,2	0,4																																				
C. vicina10	19,5	19,1	19,5	19,1	20,4	19,7	19,1	20,2	20,2	20,2	20,2	20,2	20,2	20,2	20,5	20,5	20,2	19,7	19,3	0,2	0,2	0,2	0,4	0,0	0,2	0,4	0,2																																				
C. vicina11	19,1	18,7	19,1	18,7	20,0	19,4	18,8	19,8	19,8	19,8	19,8	19,8	19,8	19,8	20,1	20,1	19,8	19,3	19,3	0,0	0,0	0,0	0,4	0,0	0,2	0,4	0,0	0,2																																				
C. vicina7	19,1	18,7	19,0	18,6	19,7	19,4	18,8	19,8	19,8	19,8	19,7	19,7	19,7	19,6	20,0	20,0	19,6	19,2	19,3	0,2	0,2	0,3	0,4	0,2	0,4	0,4	0,3	0,5	0,3																																				
C. vicina12	19,2	18,8	19,1	18,7	20,0	19,4	18,8	19,2	19,2	19,2	19,1	19,1	19,1	19,1	19,5	19,5	19,1	18,7	19,4	0,0	0,0	0,0	0,4	0,0	0,2	0,4	0,0	0,2	0,0	0,3																																				
C. vicina13	19,1	18,7	19,0	18,6	20,0	19,4	18,8	19,3	19,3	19,0	19,8	19,8	19,8	19,8	20,1	20,1	19,1	19,3	19,3	0,0	0,0	0,0	0,4	0,0	0,2	0,4	0,3	0,5	0,3	0,7	0,0																																				
C. vomitoria1	17,3	16,9	17,3	17,0	18,5	17,6	17,0	18,0	18,0	17,6	18,0	18,0	18,0	18,0	18,3	18,3	18,0	17,6	17,8	4,6	4,7	4,6	4,8	4,9	4,6	4,2	5,2	5,4	5,2	5,2	4,6	4,8																																				
C. vomitoria8	16,9	16,5	16,9	16,6	18,0	17,1	16,5	17,5	17,5	17,2	17,5	17,5	17,5	17,5	17,8	17,8	17,5	17,1	17,5	4,8	4,9	4,9	5,1	5,2	4,8	4,4	5,4	5,6	5,4	5,4	4,8	5,0	0,5																																				
C. vomitoria2	17,7	17,3	17,6	17,3	18,8	17,9	17,3	18,3	18,3	17,9	18,5	18,5	18,5	18,5	18,8	18,8	18,1	17,7	17,8	4,5	4,6	4,6	4,8	4,9	4,6	4,2	5,1	5,3	5,1	5,1	4,5	4,9	0,0	0,5																																				
C. vomitoria3	17,7	17,3	17,6	17,3	18,8	17,9	17,3	18,3	18,3	17,9	18,5	18,5	18,5	18,5	18,8	18,8	18,1	17,7	17,8	4,5	4,6	4,6	4,8	4,9	4,6	4,2	5,1	5,3	5,1	5,1	4,5	4,9	0,0	0,5	0,0																																				
C. vomitoria9	17,3	16,9	17,3	17,0	18,5	17,5	17,0	17,9	17,9	17,6	18,1	18,1	18,1	18,1	18,5	18,5	17,8	17,4	17,5	4,8	4,9	4,8	5,1	5,2	4,8	4,4	5,3	5,6	5,3	5,3	4,7	5,1	0,5	0,0	0,5	0,5																																				
C. vomitoria4	17,7	17,3	17,6	17,3	18,8	17,9	17,3	18,3	18,3	17,9	18,2	18,2	18,2	18,2	18,5	18,5	18,2	17,8	17,8	4,5	4,6	4,6	4,8	4,9	4,6	4,2	5,1	5,3	5,1	5,1	4,5	4,7	0,0	0,5	0,0	0,0	0,5																																				
C. vomitoria5	17,7	17,3	17,6	17,3	18,8	17,9	17,3	18,3	18,3	17,9	18,2	18,2	18,2	18,2	18,5	18,5	18,2	17,8	17,8	4,5	4,6	4,6	4,8	4,9	4,6	4,2	5,1	5,3	5,1	5,1	4,5	4,7	0,0	0,5	0,0	0,0	0,5	0,0																																				
C. vomitoria6	17,7	17,3	17,6	17,3	18,8	17,9	17,3	18,3	18,3	17,9	18,2	18,2	18,2	18,2	18,5	18,5	18,2	17,8	17,8	4,5	4,6	4,6	4,8	4,9	4,6	4,2	5,1	5,3	5,1	5,1	4,5	4,7	0,0	0,5	0,0	0,0	0,5	0,0	0,0																																				
C. vomitoria10	17,3	16,9	17,3	17,0	18,5	17,5	17,0	17,9	17,9	17,6	17,9	17,9	17,9	17,9	18,2	18,2	17,9	17,5	17,5	4,8	4,9	4,8	5,1	5,2	4,8	4,4	5,3	5,6	5,3	5,3	4,7	4,9	0,5	0,0	0,5	0,5	0,0	0,5	0,5	0,5																																				
C. vomitoria11	18,2	17,8	18,1	17,8	18,5	18,2	17,8	18,2	18,2	18,1	18,1	18,1	18,1	18,1	18,5	18,5	18,1	17,7	18,4	4,8	4,9	4,9	5,1	5,2	5,0	4,6	4,9	5,2	4,9	4,7	4,9	4,9	0,5	0,0	0,5	0,5	0,0	0,5	0,5	0,5	0,0																																				
C. vomitoria7	17,7	17,3	17,6	17,3	18,8	17,9	17,3	18,3	18,3	17,9	18,4	18,4	18,4	18,5	18,8	18,8	18,2	18,0	17,8	4,5	4,6	4,6	4,8	4,9	4,6	4,2	5,1	5,3	5,1	5,1	4,5	5,2	0,0	0,5	0,0	0,0	0,5	0,0	0,0	0,0	0,5	0,5																																				
C. vomitoria12	17,3	16,9	17,2	16,9	18,5	17,5	17,0	17,9	17,9	17,6	18,1	18,1	18,1	18,2	18,5	18,5	17,8	17,7	17,5	4,8	4,9	4,8	5,1	5,2	4,8	4,4	5,3	5,6	5,3	5,3	4,7	5,4	0,5	0,0	0,5	0,5	0,0	0,5	0,5	0,5	0,0	0,0	0,5																																				
L. caesar1	17,6	17,2	17,5	17,1	18,4	17,5	17,2	17,4	17,4	17,5	17,6	17,6	17,6	17,1	17,4	17,4	16,7	16,8	17,7	7,5	7,5	7,4	8,0	7,3	7,6	8,0	8,0	8,3	8,0	8,0	7,6	7,6	7,3	6,9	7,5	7,5	7,2	7,1	7,1	7,1	6,8	6,4	7,1	6,8																																				
L. caesar3	18,6	18,2	18,6	18,2	19,5	18,6	18,3	18,5	18,5	18,5	18,6	18,6	18,6	18,1	18,3	18,3	17,7	17,8	18,8	8,3	8,3	8,2	8,8	7,8	8,4	8,8	8,8	9,0	8,8	8,7	8,3	8,3	8,0	7,7	8,2	8,2	8,0	7,8	7,8	7,8	7,5	7,1	7,8	7,5	0,8																																				
L. caesar5	17,2	16,8	17,1	16,7	17,9	17,1	16,8	17,0	17,0	17,0	17,1	17,1	17,1	16,9	17,2	17,2	16,5	16,1	17,0	7,2	7,2	7,1	7,4	6,7	7,1	7,4	7,8	7,5	7,7	7,7	7,3	7,5	7,0	6,7	7,2	7,2	6,9	6,8	6,8	6,8	6,6	6,1	7,2	7,0	0,5	1,0																																				
L. caesar6	17,2	16,8	17,1	16,7	17,9	17,1	16,8	17,0	17,0	17,0	17,1	17,1	17,1	16,8	17,1	17,1	16,5	16,1	17,0	7,2	7,2	7,1	7,4	6,7	7,1	7,4	7,8	7,5	7,7	7,7	7,3	7,6	7,0	6,7	7,2	7,2	6,9	6,8	6,8	6,8	6,6	6,1	7,3	7,1	0,5	1,0	0,0																																				
L. caesar2	17,6	17,2	17,5	17,1	18,4	17,5	17,2	17,4	17,4	17,5	17,3	17,3	17,3	16,9	17,2	17,2	16,9	16,5	17,7	7,5	7,5	7,4	8,0	7,3	7,6	8,0	8,0	8,3	8,0	8,0	7,6	7,6	7,3	6,9	7,1	7,1	6,8	7,1	7,1	7,1	6,8	6,4	7,2	6,9	0,0	0,9	0,5	0,5																																				
L. caesar8	16,8	16,4	16,7	16,4	17,6	16,7	16,5	16,7	16,7	16,7	16,5	16,5	16,5	16,2	16,5	16,5	16,2	15,8	17,0	6,9	7,0	6,9	7,4	6,7	7,1	7,4	7,5	7,8	7,5	7,5	7,1	7,1	6,8	6,4	6,6	6,6	6,3	6,6	6,6	6,6	6,3	5,9	6,7	6,4	0,3	0,9	0,2	0,2	0,3																																				
L. caesar4	18,6	18,2	18,6	18,2	19,5	18,6	18,3	18,5	18,5	18,5	18,6	18,6	18,6	18,6	18,9	18,9	18,3	17,9	18,8	8,3	8,3	8,2	8,8	7,8	8,4	8,8	8,8	9,0	8,8	8,8	8,3	8,8	8,0	7,7	8,2	8,2	8,0	7,8	7,8	7,8	7,5	7,1	8,2	7,9	0,9	0,0	1,0	1,0	0,9	0,9																																				
L. caesar7	17,2	16,8	17,1	16,7	17,9	17,1	16,8	17,0	17,0	17,0	17,1	17,1	17,1	17,1	17,5	17,5	16,9	16,5	17,0	7,2	7,2	7,1	7,4	6,7	7,1	7,4	7,8	7,5	7,7	7,7	7,3	7,7	7,0	6,7	7,2	7,2	6,9	6,8	6,8	6,8	6,6	6,1	7,2	6,9	0,5	1,0	0,0	0,0	0,5	0,2	1,0																																				
H. lineatum	30,8	30,3	30,6	30,2	32,7	31,3	30,3	31,0	31,0	30,6	31,1	31,1	31,1	29,6	29,9	29,9	28,7	29,2	30,8	23,6	23,2	23,2	24,1	23,2	23,1	23,5	23,6	23,2	23,5	23,1	22,8	22,9	23,5	23,8	22,6	22,6	23,0	22,7	22,7	22,7	23,2	23,8	22,7	23,1	21,8	22,8	21,5	21,5	21,2	21,2	23,5	21,9																																				

Appendix B

(regarding to Chapter 3)

Table B1. Molecular identification of samples used in this study (68 specimens) with GenBank and BOLD databases.

Sample	Molecular identification			
	GenBank	Maximum identity (%)	BOLD	Specimen similarity (%)
An1A	Diptera sp. BOLD	99	*Delia*	100
An1B	Diptera sp. BOLD	99	*Delia*	100
An1C	Diptera sp. BOLD	100	*Delia*	100
An2	Diptera sp. BOLD	99	*Delia*	100
An3	Diptera sp. BOLD	100	*Delia*	100
An4	Diptera sp. BOLD	100	*Delia*	100
AnIII5	Diptera sp. BOLD	92	?	?
MusIII1A	*Helina evecta*	100	*Helina evecta*	100
MusIII1B	*Helina evecta*	98	*Helina evecta*	98,9
MusIII1C	*Helina evecta*	100	*Helina evecta*	100
MusIV1	*Eudasyphora cyanella*	99	*Eudasyphora cyanella*	99,8
MusIV2A	*Eudasyphora cyanella*	99	*Eudasyphora cyanella*	99,9
MusIV2B	*Eudasyphora cyanella*	100	*Eudasyphora cyanella*	100
MusIV2C	*Lucilia caesar*	100	*Lucilia caesar*	100
Hy3A	*Eudasyphora cyanella*	99	*Eudasyphora cyanella*	99,7
Hy3B	*Eudasyphora cyanella*	100	*Eudasyphora cyanella*	100
Hy3C	*Eudasyphora cyanella*	100	*Eudasyphora cyanella*	100
Hy3D	*Eudasyphora cyanella*	100	*Eudasyphora cyanella*	100
Hy3E	*Eudasyphora cyanella*	100	*Eudasyphora cyanella*	100
Hy3F	*Eudasyphora cyanella*	100	*Eudasyphora cyanella*	100
Hy3G	*Eudasyphora cyanella*	100	*Eudasyphora cyanella*	99,8
Hy3H	*Eudasyphora cyanella*	100	*Eudasyphora cyanella*	100
Hy3I	*Eudasyphora cyanella*	100	*Eudasyphora cyanella*	100
Hy3J	*Eudasyphora cyanella*	99	*Eudasyphora cyanella*	99,3
Ch1	*Lucilia illustris/Lucilia Caesar*		*Lucilia illustris/Lucilia caesar*	
Po1	*Pollenia rudis*	100	*Pollenia*	100
Po2A	*Pollenia rudis*	99	*Pollenia*	99,36
Po2B	*Pollenia rudis*	99	*Pollenia*	99,36
Po3	*Pollenia rudis*	99	*Pollenia*	99,33
Po4	*Pollenia rudis*	100	*Pollenia*	100
Po5A	*Pollenia rudis*	100	*Pollenia*	100
Po5B	Diptera sp. BOLD	90	?	?
Po5C	*Pollenia rudis*	99	*Pollenia*	99,17
Po5E	*Pollenia rudis*	99	*Pollenia*	100

Table B1 (cont.). Molecular identification of samples used in this study (68 specimens) with GenBank and BOLD databases.

Sample	Molecular identification			
	GenBank	Maximum identity (%)	BOLD	Specimen similarity (%)
Ph1	*Phaonia subventa*	99	*Phaonia subventa*	99,17
Ph2A	Diptera sp. BOLD	90	?	?
Ph2B	Diptera sp. BOLD	90	?	?
Ph2C	Diptera sp. BOLD	90	?	?
Ph3A	Diptera sp. BOLD	93	*Phaonia errans*	93,94
Ph3B	Diptera sp. BOLD	96	?	?
Ph4	*Phaonia subventa*	99	*Phaonia subventa*	99,12
Ph5	*Phaonia subventa*	99	*Phaonia subventa*	99,1
MuI1	*Musca autumnalis*	99	*Musca autumnalis*	99,7
MuI3	*Musca autumnalis*	99	*Musca autumnalis*	99,7
MuI4	*Musca autumnalis*	99	*Musca autumnalis*	100
MuI5	*Musca autumnalis*	99	*Musca autumnalis*	99,7
MuII7	*Muscina levida/Muscina assimilis*		*Muscina levida/Muscina assimilis*	
MuII8	*Muscina levida/Muscina assimilis*		*Muscina levida/Muscina assimilis*	
MuII9A	Diptera sp. BOLD	98	*Helina reversio*	99
MuII9B	Diptera sp. BOLD	98	*Helina*	98
MuII9C	Diptera sp. BOLD	98	*Helina reversio*	99
MuIII10	*Phaonia tugoriorum*	99	*Phaonia tuguriorum*	100
MuIII11	Diptera sp. BOLD	90	?	?
MuIII12	*Helina impucta*	98	*Helina*	98
MuIII13A	*Hydrotaea dentipes*	99	*Hydrotaea dentipes*	100
MuIII13B	*Hydrotaea dentipes*	99	*Hydrotaea dentipes*	100
MuIII14A	*Phaonia tugoriorum*	100	*Phaonia tuguriorum*	100
MuIII14C	*Phaonia tuguriorum*	100	*Phaonia tuguriorum*	100
Hy1B	*Hydrotaea dentipes*	100	*Hydrotaea dentipes*	100
Hy1C	*Hydrotaea dentipes*	100	*Hydrotaea dentipes*	100
Hy1D	*Hydrotaea dentipes*	99	*Hydrotaea dentipes*	100
Hy1E	*Hydrotaea dentipes*	100	*Hydrotaea dentipes*	100
Hy1F	Diptera sp. BOLD	93	*Hydrotaea armipes*	100
Hy2A	*Hydrotaea dentipes*	99	*Hydrotaea dentipes*	100
Hy2B	*Hydrotaea dentipes*	100	*Hydrotaea dentipes*	100
Hy2C	*Hydrotaea dentipes*	100	*Hydrotaea dentipes*	100
Hy2D	*Hydrotaea dentipes*	100	*Hydrotaea dentipes*	100
Hy2E	Diptera sp. BOLD	100	*Muscina*	100

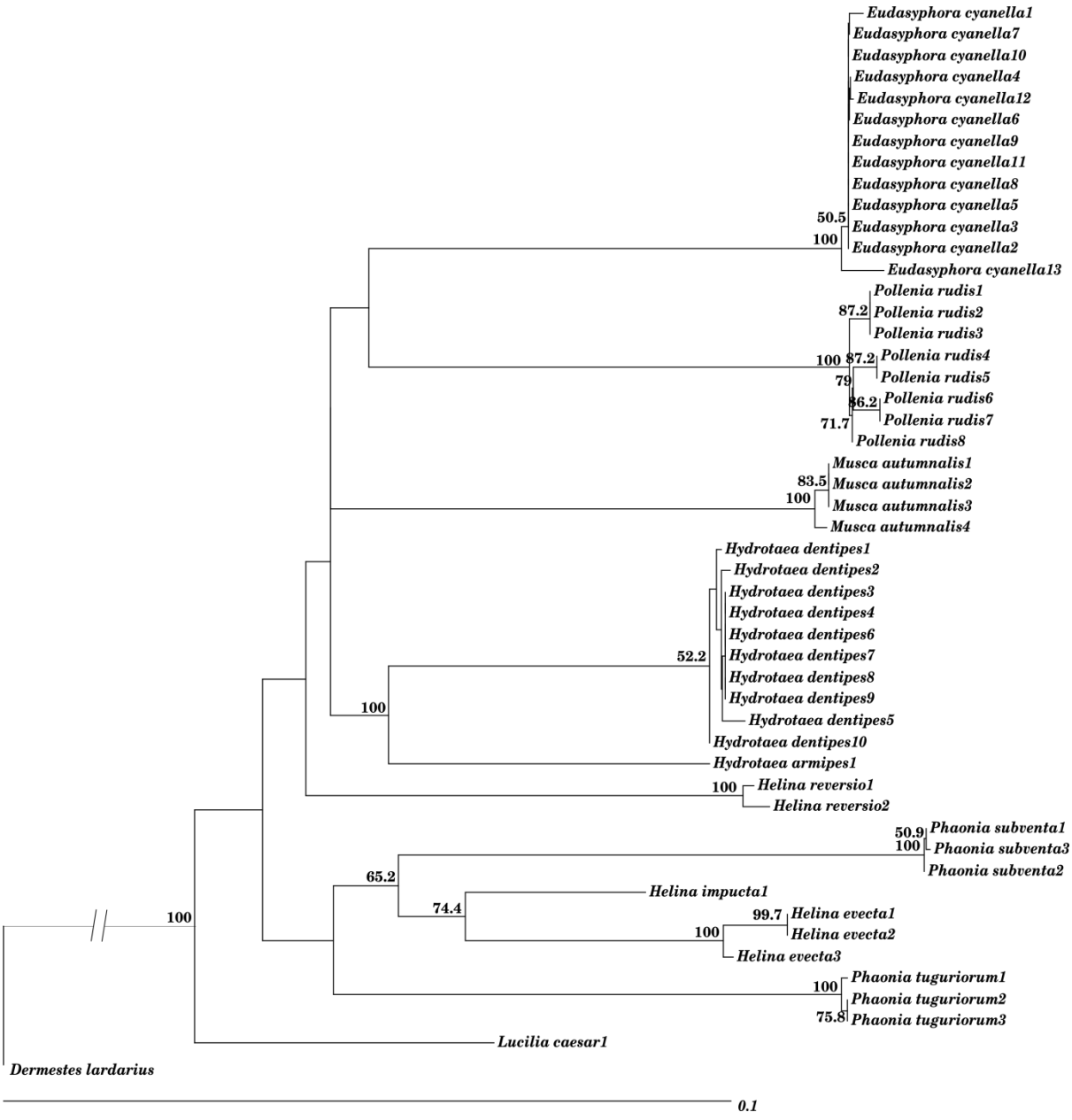

Figure B1. Neighbor-joining phylogram of 69 cytochrome *c* oxidase I (COI) sequences from ten Diptera species and one outgroup (*Dermestes lardarius*). Bootstrap values indicate support for nodes among 1000 bootstrap replicates.

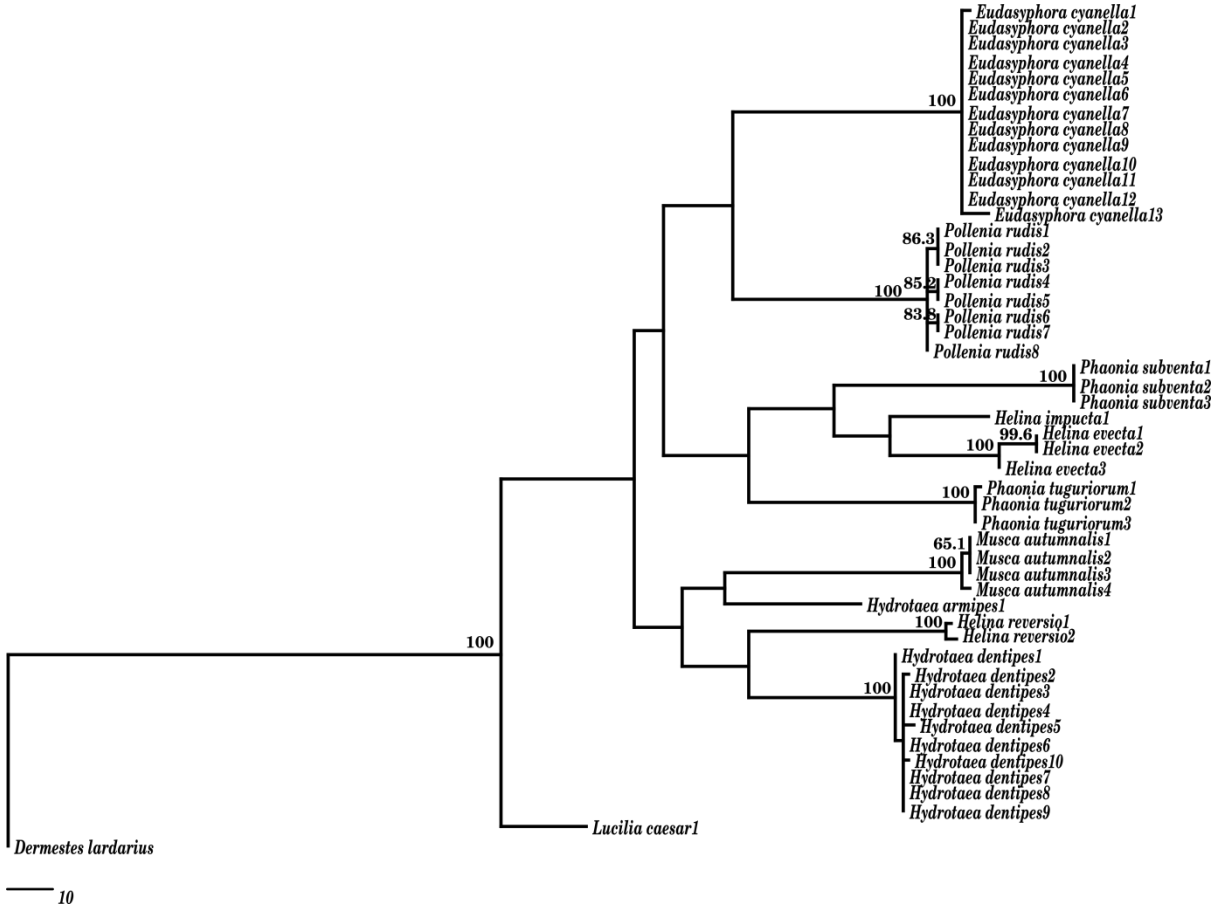

Figure B2. Maximum parsimony phylogram of heuristic search procedure (Tree Bisection and Reconnection algorithm, TBR) for 69 cytochrome *c* oxidase I (COI) sequences from ten Diptera species and one outgroup (*Dermestes lardarius*). Bootstrap values indicate support for nodes among 1000 bootstrap replicates (heuristic search procedure).

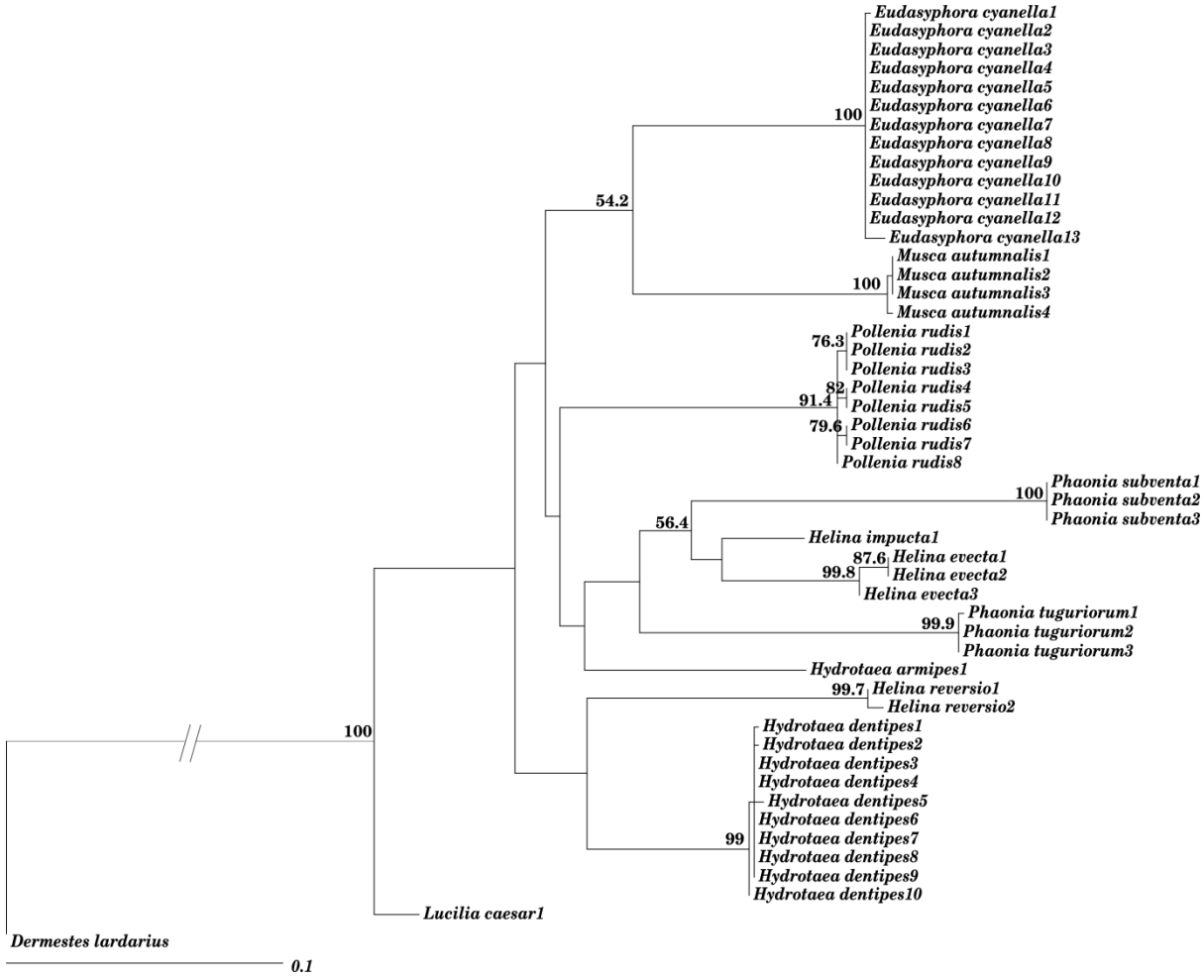

Figure B3. Maximum likelihood phylogram of 53 cytochrome *c* oxidase I (COI) sequences from ten Diptera species and one outgroup (*Dermestes lardarius*). Bootstrap values indicate support for nodes among 1000 bootstrap replicates.

Figure B4. Bayesian phylogeny of 69 cytochrome *c* oxidase I (COI) sequences from ten Diptera species and one outgroup (*Hypoderma lineatum*). Values on tree branches indicate posterior probability for nodes.

Table B2. Percentage of nucleotide sequence divergence (p-distances) at cytochrome *c* oxidase I (COI) region for ten Diptera species. E = *Eudasyphora*; L = *Lucilia*; Po = *Pollenia*; M = *Musca*; Ph = *Phaonia*; He = *Helina*; Hy = *Hydrotaea*.

	E. cyanella1	*E. cyanella2*	*E. cyanella3*	*E. cyanella4*	*E. cyanella5*	*E. cyanella6*	*E. cyanella7*	*E. cyanella8*	*E. cyanella9*	*E. cyanella10*	*E. cyanella11*	*E. cyanella12*	*E. cyanella13*	*L. caesar1*	*Po. rudis1*	*Po. rudis2*	*Po. rudis3*	*Po. rudis4*	*Po. rudis5*	*Po. rudis6*	*Po. rudis7*	*Po. rudis8*	*M. autumnalis1*	*M. autumnalis2*	*M. autumnalis4*	*M. autumnalis3*	*Ph. subventa1*	*Ph. subventa2*	*Ph. subventa3*	*Ph. tuguriorum1*	*Ph. tuguriorum2*	*Ph. tuguriorum3*	*He. impucta1*	*He. evecta1*	*He. evecta3*	*He. evecta2*	*He. reversio1*	*He. reversio2*	*Hy. dentipes1*	*Hy. dentipes2*	*Hy. dentipes3*	*Hy. dentipes4*	*Hy. dentipes5*	*Hy. dentipes6*	*Hy. dentipes10*	*Hy. dentipes7*	*Hy. dentipes8*	*Hy. dentipes9*	*Hy. armipes1*
E. cyanella1	0,5	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	1,0	11,9	12,5	12,5	12,5	12,7	13,4	12,5	12,6	12,5	12,0	11,9	11,9	11,9	15,8	15,9	15,8	15,1	14,7	15,1	12,0	14,7	14,5	14,5	12,5	12,8	11,9	12,0	11,5	11,9	12,0	11,9	11,1	11,9	11,5	11,9	11,9
E. cyanella2		0,2	0,2	0,0	0,2	0,2	0,2	0,0	0,0	0,0	0,0	1,0	11,6	12,5	12,5	12,5	12,6	13,2	12,5	12,4	12,1	12,0	11,9	11,9	11,9	15,7	15,7	15,6	14,9	14,5	14,9	11,9	14,4	14,5	14,5	12,5	12,6	11,7	11,9	11,5	11,7	11,8	11,7	10,9	11,7	11,5	11,7	11,9
E. cyanella3			0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,8	11,7	12,0	12,1	12,1	12,5	12,5	12,0	12,4	12,1	12,1	12,0	11,9	12,0	15,9	15,7	15,5	14,5	14,5	14,5	11,7	14,5	15,7	14,1	12,5	12,4	11,5	11,7	11,1	11,5	11,7	11,5	10,8	11,5	11,1	11,5	11,6
E. cyanella4				0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,8	11,7	12,5	12,5	12,5	12,5	12,9	12,5	12,4	12,1	12,0	11,9	11,9	11,9	15,8	15,7	15,6	14,9	14,5	14,8	11,9	14,5	14,1	14,5	12,5	12,5	11,6	11,7	11,2	11,6	11,8	11,6	10,9	11,6	11,2	11,6	11,6
E. cyanella5					0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,8	12,0	12,7	12,7	12,7	12,4	12,6	12,7	12,4	12,1	12,0	12,0	12,0	12,0	15,7	15,6	15,6	14,8	14,4	14,8	11,9	14,5	14,0	14,5	12,5	12,5	11,2	11,4	11,2	11,2	11,5	11,2	10,9	11,2	11,2	11,2	11,6
E. cyanella6						0,0	0,0	0,0	0,0	0,0	0,0	0,8	11,9	12,5	12,5	12,5	12,5	12,9	12,5	12,4	12,1	11,9	11,9	11,9	11,9	15,9	15,7	15,6	14,9	14,6	14,9	11,9	14,6	14,2	14,5	12,5	12,5	11,5	11,4	11,2	11,5	11,4	11,5	10,9	11,5	11,2	11,5	11,5	11,5
E. cyanella7							0,0	0,0	0,0	0,0	0,0	0,8	11,8	12,4	12,4	12,4	12,5	12,9	12,4	12,4	12,1	12,0	11,8	11,8	11,8	15,7	15,7	15,6	14,8	14,4	14,8	11,8	14,5	14,0	14,5	12,2	12,5	11,5	11,6	11,1	11,5	11,5	11,5	10,9	11,5	11,2	11,5	11,5	11,5
E. cyanella8								0,0	0,0	0,0	0,0	0,8	12,0	12,5	12,5	12,5	12,5	12,5	12,5	12,4	12,1	12,1	12,0	12,0	12,0	15,9	15,6	15,6	14,5	14,5	14,5	11,7	14,5	15,5	14,1	12,5	12,4	11,2	11,5	11,1	11,2	11,4	11,2	10,9	11,2	11,1	11,2	11,5
E. cyanella9									0,0	0,0	0,0	0,7	11,8	12,1	12,5	12,5	12,5	12,1	12,1	12,2	11,9	11,9	12,0	11,9	12,0	15,7	15,6	15,6	14,5	14,2	14,5	11,5	14,1	15,4	14,0	12,1	12,5	10,9	11,1	10,9	10,9	11,5	10,9	10,7	10,9	11,0	10,9	11,5
E. cyanella10										0,0	0,0	0,7	11,8	12,1	12,5	12,5	12,5	12,1	12,1	12,2	11,9	11,9	12,0	11,9	12,0	15,7	15,6	15,6	14,5	14,2	14,5	11,5	14,1	15,4	14,0	12,1	12,5	10,9	11,1	10,9	10,9	11,5	10,9	10,7	10,9	11,0	10,9	11,5
E. cyanella11											0,0	0,7	11,8	12,1	12,5	12,5	12,5	12,1	12,1	12,2	11,9	11,9	12,0	11,9	12,0	15,7	15,6	15,6	14,5	14,2	14,5	11,5	14,1	15,4	14,0	12,1	12,5	10,9	11,1	10,9	10,9	11,5	10,9	10,7	10,9	11,0	10,9	11,5
E. cyanella12												0,7	11,4	12,5	12,5	12,5	12,5	12,5	12,5	12,4	12,1	11,8	11,8	11,8	11,8	15,8	15,7	15,7	14,6	14,5	14,6	11,4	14,2	15,5	14,0	12,5	12,5	11,1	11,2	11,1	11,1	11,5	11,1	10,8	11,1	11,1	11,1	11,2
E. cyanella13													12,2	12,2	12,4	12,4	12,6	12,5	12,2	12,5	12,2	12,4	12,2	12,2	12,2	16,1	16,0	15,9	14,8	14,7	14,8	11,9	14,8	14,0	14,4	12,7	12,9	11,4	11,6	11,2	11,4	11,8	11,4	11,0	11,4	11,2	11,4	11,9
L. caesar1														10,9	10,8	10,8	11,4	11,7	10,9	11,5	10,9	10,7	10,8	11,0	10,8	12,5	12,8	12,5	15,8	15,6	15,7	9,4	11,7	11,5	11,5	11,5	11,5	10,8	10,8	10,4	10,9	11,2	10,9	10,2	10,9	10,4	10,9	10,5
Po. rudis1															0,8	0,8	0,7	0,5	0,0	0,8	0,5	15,4	15,4	15,5	15,4	14,5	14,2	14,4	15,7	15,4	15,5	11,1	15,4	12,6	15,0	12,9	12,8	12,5	12,6	12,1	12,5	12,9	12,5	12,2	12,5	12,1	12,5	12,2
Po. rudis2																0,0	0,7	1,1	0,8	0,8	0,5	15,7	15,7	15,6	15,7	14,1	14,2	14,4	15,2	12,9	15,1	11,5	15,4	12,8	15,1	12,9	12,6	12,5	12,8	12,5	12,6	12,9	12,6	12,2	12,6	12,5	12,6	12,5
Po. rudis3																	0,7	1,1	0,8	0,8	0,5	15,7	15,7	15,6	15,7	14,1	14,2	14,4	15,2	12,9	15,1	11,5	15,4	12,8	15,1	12,9	12,6	12,5	12,8	12,5	12,6	12,9	12,6	12,2	12,6	12,5	12,6	12,5
Po. rudis4																		0,7	0,7	0,0	0,5	15,9	15,9	15,5	15,9	14,2	14,2	14,4	15,7	15,5	15,5	11,4	15,0	12,5	12,9	12,5	12,6	12,0	12,4	12,5	12,2	12,5	12,2	12,1	12,2	12,5	12,2	11,9
Po. rudis5																			0,5	0,8	0,5	14,0	14,0	15,6	14,0	14,7	14,2	14,4	14,5	15,5	14,2	11,6	15,0	12,7	15,2	12,9	15,2	12,4	12,7	12,6	12,5	12,5	12,5	12,2	12,5	12,6	12,5	12,4
Po. rudis6																				0,8	0,5	15,4	15,4	15,5	15,4	14,5	14,2	14,4	15,7	15,4	15,5	11,1	15,4	12,6	15,0	12,9	12,8	12,5	12,6	12,1	12,5	12,9	12,5	12,2	12,5	12,1	12,5	12,2
Po. rudis7																					0,5	15,7	15,7	15,4	15,7	14,5	14,5	14,4	15,9	15,7	15,7	11,4	12,9	12,4	12,9	12,6	12,5	12,1	12,4	12,5	12,5	12,5	12,5	12,0	12,5	12,5	12,5	11,8
Po. rudis8																						15,4	15,5	15,1	15,5	14,0	15,9	14,1	15,5	15,2	15,5	11,1	12,6	11,9	12,6	12,6	12,6	11,8	12,1	11,9	11,9	12,2	11,9	11,7	11,9	12,0	11,9	11,6
M. autumnalis1																							0,0	0,5	0,0	15,9	14,1	15,9	15,4	15,4	15,4	12,0	15,9	15,2	15,8	15,8	15,9	11,1	11,0	11,0	11,1	11,0	11,1	10,5	11,1	11,0	11,1	10,7
M. autumnalis2																								0,5	0,0	14,1	14,0	15,9	15,5	15,4	15,5	12,0	15,9	15,2	15,8	15,7	15,9	11,4	11,5	11,0	11,4	11,5	11,4	10,5	11,4	11,0	11,4	10,8
M. autumnalis4																									0,5	14,4	14,1	15,9	15,6	15,4	15,6	12,2	15,9	15,4	15,8	15,7	14,0	11,7	11,6	11,5	11,7	11,4	11,7	10,7	11,7	11,4	11,7	11,0
M. autumnalis3																										14,1	14,0	15,9	15,5	15,4	15,5	12,0	15,9	15,2	15,8	15,7	15,9	11,4	11,5	11,0	11,4	11,5	11,4	10,5	11,4	11,0	11,4	10,8
Ph. subventa1																											0,0	0,0	14,6	14,5	14,7	10,2	11,5	10,9	11,2	15,7	14,1	15,5	15,2	15,1	15,4	15,8	15,4	15,2	15,4	15,1	15,4	12,9
Ph. subventa2																												0,0	14,2	14,4	14,4	10,4	11,5	10,6	11,1	15,9	14,5	15,6	15,5	15,4	15,4	15,5	15,4	15,5	15,4	15,4	15,4	12,9
Ph. subventa3																													14,6	14,7	14,8	10,5	11,2	10,5	11,2	14,1	14,6	15,7	15,4	15,6	15,6	15,7	15,6	15,4	15,6	15,6	15,6	12,9
Ph. tuguriorum1																																0,2	0,2	10,8	10,9	11,4	11,5	12,5	12,9	15,7	15,7	15,4	15,5	15,1	15,6	12,8	15,5	15,4	15,5	15,7
Ph. tuguriorum2																																	0,0	10,4	11,0	10,8	11,0	12,1	12,4	15,1	15,4	15,1	15,5	15,2	15,5	12,9	15,5	15,1	15,5	15,5
Ph. tuguriorum3																																		10,6	11,0	11,5	11,5	12,2	12,8	15,6	15,8	15,6	15,7	15,5	15,7	12,9	15,7	15,6	15,7	15,6
He. impucta1																																			6,7	6,1	6,6	10,9	11,0	11,5	11,5	11,5	11,4	12,1	11,4	11,4	11,4	11,5	11,4	9,6
He. evecta1																																				1,1	0,0	12,4	12,8	12,7	12,4	12,4	12,6	12,7	12,6	12,2	12,6	12,4	12,6	11,2
He. evecta3																																					1,2	12,0	12,5	11,9	11,9	11,8	12,0	12,4	12,0	12,1	12,0	11,8	12,0	10,5
He. evecta2																																						12,6	12,8	12,5	12,2	12,5	12,5	12,5	12,5	12,2	12,5	12,5	12,5	10,9
He. reversio1																																							0,5	9,7	10,0	9,7	9,9	10,1	9,9	9,5	9,9	9,7	9,9	11,5
He. reversio2																																								9,9	10,2	10,1	10,0	10,5	10,0	9,7	10,0	10,1	10,0	11,5
Hy. dentipes1																																									0,5	0,2	0,2	0,5	0,2	0,5	0,2	0,2	0,2	7,9
Hy. dentipes2																																										0,2	0,2	0,5	0,2	0,5	0,2	0,2	0,2	7,9
Hy. dentipes3																																											0,0	0,4	0,0	0,2	0,0	0,0	0,0	8,1
Hy. dentipes4																																												0,5	0,0	0,2	0,0	0,0	0,0	8,1
Hy. dentipes5																																													0,5	0,5	0,5	0,4	0,5	8,5
Hy. dentipes6																																														0,2	0,0	0,0	0,0	8,1
Hy. dentipes10																																															0,2	0,2	0,2	8,2
Hy. dentipes7																																																0,0	0,0	8,1
Hy. dentipes8																																																0,0	8,1
Hy. dentipes9																																																	8,1
Hy. armipes1																																																	

Table B2. Percentage of nucleotide sequence divergence (ML distances) at cytochrome c oxidase I (COI) region for ten Diptera species. E = *Eudasyphora*; L = *Lucilia*; Po = *Pollenia*; M = *Musca*; Ph = *Phaonia*; He = *Helina*; Hy = *Hydrotaea*.

	E. cyanella1	*E. cyanella2*	*E. cyanella3*	*E. cyanella4*	*E. cyanella5*	*E. cyanella6*	*E. cyanella7*	*E. cyanella8*	*E. cyanella9*	*E. cyanella10*	*E. cyanella11*	*E. cyanella12*	*E. cyanella13*	*L. caesar1*	*Po. rudis1*	*Po. rudis2*	*Po. rudis3*	*Po. rudis4*	*Po. rudis5*	*Po. rudis6*	*Po. rudis7*	*Po. rudis8*	*M. autumnalis1*	*M. autumnalis2*	*M. autumnalis4*	*M. autumnalis3*	*Ph. subventa1*	*Ph. subventa2*	*Ph. subventa3*	*Ph. tuguriorum1*	*Ph. tuguriorum2*	*Ph. tuguriorum3*	*He. impucta1*	*He. evecta1*	*He. evecta3*	*He. evecta2*	*He. reversio1*	*He. reversio2*	*Hy. dentipes1*	*Hy. dentipes2*	*Hy. dentipes3*	*Hy. dentipes4*	*Hy. dentipes5*	*Hy. dentipes6*	*Hy. dentipes10*	*Hy. dentipes7*	*Hy. dentipes8*	*Hy. dentipes9*	*Hy. armipes1*
E. cyanella1																																																				
E. cyanella2	0,3																																																			
E. cyanella3	0,2	0,2																																																		
E. cyanella4	0,2	0,2	0,0																																																
E. cyanella5	0,2	0,0	0,0	0,0																																															
E. cyanella6	0,2	0,2	0,0	0,0	0,0																																														
E. cyanella7	0,2	0,2	0,0	0,0	0,0	0,0																																													
E. cyanella8	0,2	0,2	0,0	0,0	0,0	0,0	0,0																																												
E. cyanella9	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0																																											
E. cyanella10	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0																																									
E. cyanella11	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0																																							
E. cyanella12	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0																																						
E. cyanella13	1,0	1,0	0,8	0,8	0,8	0,8	0,8	0,8	0,7	0,7	0,7	0,7																																					
L. caesar1	13,2	12,9	13,0	13,0	13,8	13,2	13,1	13,4	13,1	13,1	13,1	12,7	13,6																																				
Po. rudis1	13,9	13,8	13,8	13,7	14,1	13,9	13,8	13,7	13,4	13,4	13,4	13,8	13,6	12,1																																			
Po. rudis2	13,9	13,7	13,8	13,7	14,1	13,9	13,8	13,9	13,6	13,6	13,6	14,0	13,7	11,9	0,8																																		
Po. rudis3	13,9	13,7	13,8	13,7	14,1	13,9	13,8	13,9	13,6	13,6	13,6	14,0	13,7	11,9	0,8	0,0																																	
Po. rudis4	14,1	14,0	14,0	14,0	13,8	14,0	14,0	14,0	13,6	13,6	13,6	13,9	14,2	12,8	0,7	0,7	0,7																																
Po. rudis5	14,9	14,8	13,7	14,4	14,1	14,4	14,4	13,8	13,4	13,4	13,4	13,8	13,9	12,9	0,8	1,1	1,1	0,7																															
Po. rudis6	13,9	13,8	13,8	13,7	14,1	13,9	13,8	13,7	13,4	13,4	13,4	13,8	13,6	12,1	0,0	0,8	0,8	0,7	0,8																														
Po. rudis7	13,9	13,7	13,8	13,7	13,7	13,7	13,7	13,8	13,6	13,6	13,6	13,9	14,0	12,8	0,8	0,8	0,8	0,0	0,8	0,8																													
Po. rudis8	13,7	13,6	13,6	13,6	13,6	13,6	13,6	13,6	13,8	13,8	13,8	13,6	13,7	12,0	0,8	0,8	0,8	0,8	0,8	0,8	0,8																												
M. autumnalis1	13,2	13,2	13,8	13,1	13,2	13,1	13,1	13,8	13,1	13,1	13,1	13,0	13,7	11,8	15,0	15,4	15,4	15,6	15,8	15,0	15,3	15,1																												
M. autumnalis2	13,1	13,1	13,2	13,0	13,2	13,0	13,0	13,2	13,1	13,1	13,1	13,0	13,6	12,0	15,0	15,4	15,4	15,6	15,7	15,0	15,3	15,1	0,0																											
M. autumnalis4	13,1	13,1	13,2	13,0	13,2	13,0	13,0	13,2	13,1	13,1	13,1	13,0	13,6	12,1	14,8	15,2	15,2	15,1	15,8	14,8	14,9	14,7	0,3	0,3																										
M. autumnalis3	13,1	13,1	13,2	13,0	13,2	13,0	13,0	13,2	13,1	13,1	13,1	13,0	13,6	12,0	15,0	15,4	15,4	15,6	15,7	15,0	15,3	15,1	0,0	0,0	0,3																									
Ph. subventa1	18,6	18,4	18,7	18,5	18,1	18,5	18,4	18,5	18,2	18,2	18,2	18,6	18,8	14,3	16,5	16,3	16,3	16,3	16,8	16,5	16,2	16,0	16,0	16,4	16,7	16,4																								
Ph. subventa2	18,6	18,4	18,4	18,4	18,2	18,4	18,4	18,3	18,2	18,2	18,2	18,6	18,8	14,6	16,4	16,4	16,4	16,3	16,3	16,4	16,4	16,1	16,3	16,3	16,3	16,3	0,0																							
Ph. subventa3	18,4	18,2	18,2	18,2	18,2	18,2	18,2	18,2	18,2	18,2	18,2	18,6	18,6	14,4	16,6	16,6	16,6	16,6	16,6	16,6	16,6	16,2	16,1	16,1	16,1	16,1	0,0	0,0																						
Ph. tuguriorum1	17,2	17,0	16,5	16,9	16,8	16,9	16,8	16,5	16,2	16,2	16,2	16,7	16,8	15,8	15,4	14,8	14,8	15,3	16,1	15,4	15,5	15,1	17,7	17,8	17,9	17,8	16,9	16,4	16,9																					
Ph. tuguriorum2	16,7	16,5	16,5	16,5	16,3	16,5	16,4	16,5	16,2	16,2	16,2	16,7	16,8	15,4	15,0	14,4	14,4	15,1	15,1	15,0	15,3	14,9	17,6	17,7	17,7	17,7	16,7	16,6	17,1	0,2																				
Ph. tuguriorum3	17,2	17,0	16,5	16,9	16,8	16,9	16,8	16,5	16,2	16,2	16,2	16,7	16,8	15,6	15,2	14,6	14,6	15,1	15,9	15,2	15,3	14,9	17,7	17,8	17,9	17,8	17,1	16,6	17,1	0,2	0,0																		
He. impucta1	13,5	13,5	13,0	13,5	13,2	13,5	13,2	13,1	12,8	12,8	12,8	12,8	13,3	10,4	12,4	12,5	12,5	12,6	12,8	12,4	12,6	12,3	13,4	13,5	13,7	13,5	11,4	11,7	11,8	11,9	11,4	11,7																	
He. evecta1	16,9	16,5	16,7	16,7	16,4	16,7	16,6	16,5	16,2	16,2	16,2	16,4	17,0	13,1	15,2	15,2	15,2	14,7	14,7	15,2	14,6	14,2	15,9	15,8	15,8	15,8	13,0	12,7	12,7	12,1	12,3	12,3	7,2															
He. evecta3	16,4	16,4	15,7	16,1	15,9	16,1	16,0	15,4	15,3	15,3	15,3	15,5	16,0	12,7	14,3	14,5	14,5	13,9	14,3	14,3	13,9	13,4	15,1	15,1	15,3	15,1	12,4	11,9	11,8	12,8	11,9	12,6	6,5	1,1													
He. evecta2	16,6	16,4	16,2	16,4	16,4	16,4	16,4	16,3	16,1	16,1	16,1	16,3	16,6	13,0	14,7	14,9	14,9	14,7	14,9	14,7	14,5	14,3	15,8	15,8	15,8	15,8	12,6	12,6	12,7	12,7	12,2	12,9	7,1	0,0	1,2												
He. reversio1	14,0	13,8	13,8	13,8	13,8	13,8	13,7	13,9	13,6	13,6	13,6	14,1	14,3	12,9	14,6	14,6	14,6	14,0	14,5	14,6	14,1	14,2	15,7	15,7	15,7	15,7	15,9	16,0	16,3	13,8	13,6	13,6	12,2	14,2	13,6	14,5											
He. reversio2	14,4	14,2	13,9	14,0	14,0	14,0	14,0	14,0	13,8	13,8	13,8	14,1	14,5	12,7	14,4	14,2	14,2	14,1	14,9	14,4	14,0	14,2	15,9	15,9	15,9	15,9	16,2	16,8	16,9	14,6	13,9	14,4	12,3	14,6	14,2	14,7	0,5									
Hy. dentipes1	13,2	13,0	12,7	12,8	12,8	12,4	12,5	12,3	12,0	12,0	12,0	12,8	12,6	12,1	13,8	14,0	14,0	13,4	13,8	13,8	13,8	13,1	12,5	12,8	13,1	12,8	15,6	15,6	15,8	15,5	14,7	15,3	12,7	14,6	13,5	14,2	10,8	11,0								
Hy. dentipes2	13,4	13,2	12,9	13,0	12,5	12,6	12,7	12,6	12,2	12,2	12,2	12,5	12,8	12,1	14,2	14,4	14,4	13,8	14,2	14,2	13,7	13,5	12,3	12,6	13,0	12,6	15,2	15,2	15,4	15,5	15,1	15,7	12,7	14,2	13,5	13,8	11,1	11,3	0,8							
Hy. dentipes3	12,7	12,5	12,2	12,3	12,3	12,3	12,3	12,2	12,0	12,0	12,0	12,3	12,4	11,5	13,6	13,7	13,7	13,6	14,1	13,6	13,5	13,3	12,3	12,3	12,6	12,3	15,0	15,4	15,6	15,1	14,7	15,3	12,7	14,1	13,4	14,0	10,7	11,1	0,2	0,2					
Hy. dentipes4	13,2	13,0	12,7	12,8	12,8	12,4	12,5	12,3	12,0	12,0	12,0	12,8	12,6	12,3	14,0	14,2	14,2	13,6	14,0	14,0	13,5	13,3	12,5	12,8	13,1	12,8	15,4	15,4	15,6	15,3	14,9	15,5	12,9	14,3	13,6	14,0	11,0	11,1	0,2	0,2	0,0				
Hy. dentipes5	13,3	13,1	13,1	13,1	12,6	12,7	12,8	12,7	12,6	12,6	12,6	12,5	13,2	12,7	14,5	14,5	14,5	13,8	14,0	14,5	13,9	13,7	12,3	12,7	12,9	12,7	16,2	15,7	16,0	14,8	15,0	15,0	13,7	14,6	14,2	14,4	11,3	11,5	0,5	0,5	0,4	0,3			
Hy. dentipes6	13,2	13,0	12,7	12,8	12,8	12,4	12,5	12,3	12,0	12,0	12,0	12,8	12,6	12,3	14,0	14,2	14,2	13,6	14,0	14,0	13,5	13,3	12,5	12,8	13,1	12,8	15,4	15,4	15,6	15,3	14,9	15,5	12,9	14,3	13,6	14,0	11,0	11,1	0,2	0,2	0,0	0,0	0,3		
Hy. dentipes10	12,2	12,0	12,0	12,0	12,0	12,0	12,0	12,1	11,9	11,9	11,9	12,1	12,2	11,3	13,6	13,6	13,6	13,5	13,6	13,6	13,4	13,1	11,7	11,7	11,9	11,7	15,1	15,3	15,4	14,3	14,5	14,5	12,8	13,9	13,7	13,9	10,6	10,7	0,8	0,8	0,2	0,2	0,5	0,2	
Hy. dentipes7	13,2	13,0	12,7	12,8	12,8	12,4	12,5	12,3	12,0	12,0	12,0	12,8	12,6	12,3	14,0	14,2	14,2	13,6	14,0	14,0	13,5	13,3	12,5	12,8	13,1	12,8	15,4	15,4	15,6	15,3	14,9	15,5	12,9	14,3	13,6	14,0	11,0	11,1	0,2	0,2	0,0	0,0	0,3	0,0	0,2
Hy. dentipes8	12,8	12,6	12,2	12,3	12,3	12,3	12,3	12,3	12,1	12,1	12,1	12,3	12,4	11,6	13,6	13,8	13,8	13,7	14,1	13,6	13,6	13,3	12,3	12,3	12,7	12,3	15,0	15,5	15,6	15,2	14,7	15,4	12,8	14,1	13,4	14,0	10,8	11,2	0,2	0,2	0,0	0,0	0,4	0,0	0,2	0,0
Hy. dentipes9	13,2	13,0	12,7	12,8	12,8	12,4	12,5	12,3	12,0	12,0	12,0	12,3	12,6	12,3	14,0	14,2	14,2	13,6	14,0	14,0	13,5	13,3	12,5	12,8	13,1	12,8	15,4	15,4	15,6	15,3	14,9	15,5	12,9	14,3	13,6	14,0	11,0	11,1	0,2	0,2	0,0	0,0	0,3	0,0	0,2	0,0	0,0
Hy. armipes1	13,3	13,3	13,0	12,9	12,8	12,5	12,6	12,9	12,8	12,8	12,8	12,6	13,3	11,5	13,7	13,9	13,9	13,2	13,8	13,7	13,0	12,7	11,9	12,0	12,2	12,0	14,9	14,9	14,8	15,6	15,0	15,4	10,6	12,5	11,7	12,1	12,8	12,8	8,6	8,6	8,9	8,8	9,1	8,8	9,0	8,8	8,9	8,8